ASSORTATIVE MARRIAGE IN EDINBURGH

GRAHAM SUTTON

A thesis presented for the degree of
Doctor of Philosophy of the
University of Edinburgh in the
Faculty of Medicine.

September 1980

CONTENTS

S U M M A R Y

Assortative marriage means the tendency for people to choose spouses who resemble each other more closely than one would expect by chance. It may also promote social diversity, and govern the "inheritance" of social traits such as religion and voting patterns which are not genetic but are strongly familial.

A study of assortative marriage was undertaken in a series of Edinburgh couples, in order to determine

a) the strength of assortative marriage for a wide range of physical, psychological and social traits.

b) to what extent these similarities were present at the time of marriage, and whether the couples had converged or diverged since marriage.

c) the nature of the interaction of those traits which were interlinked.

d) the implications of these similarities for those genetic and social traits which assortative marriage is believed to influence.

By confining the study to those traits which can be quickly and objectively measured, it was possible to study a wide variety of traits and yet have a large sample size. Eventually 68 engaged couples, 113 newly weds, and 222 couples who had been married for five years or more were ascertained and visited at home.

As many as possible of these 403 couples were then revisited a year later, to remeasure the traits. A computer program was written to handle the large volume of data collected.

General traits measured included age at marriage, fertility of the couple, number of sibs of each partner, and their previous marital status. Age was highly correlated, as with other studies. The correlation was higher in remarriages than first marriages, and lower in old than young couples.

Fertility, i.e. total planned family size was of course strongly correlated; there was also evidence of a marked reduction in plans for total planned family size as the marriage progressed. Family plans were not related to education or social class.

There was little or no association for number of sibs. Previous marital status was highly correlated even after correction for age. In young couples, bachelor with spinster and divorcee with divorcee were preferred combinations, and widows/widowers showed no preference. This was probably related to previous fertility. In older couples, widow with widower was preferred but there was no distinction between singletons and divorcees.

Physical traits measured included height, weight, skinfold thickness, blood pressure, and pulse before and after exercise. Height was correlated even after taking account of age and this leads to an increase in the population variance for height. However there is a complex interaction with perceptions of social class. Weight, assessed in a number of ways, did not show association in this sample. The various cardiovascular traits occasionally showed correlations but these all proved to be artefacts of age, smoking etc.

Psychological traits studied were neuroticism and extroversion, as seen in oneself and in one's partner and measured by Eysenck's Personality Inventory. Couples showed assortative marriage for neuroticism; curiously they assorted for extroversion only in newly weds. It was expected that these psychological traits might show convergence during the course of the marriage. In fact there was no evidence of this; the results for extroversion may even have indicated divergence.

Social traits examined were occupation, father's occupation, smoking habits, education, and religious practice. There was assortative marriage for all these traits but all were highly intercorrelated. Initially occupation appeared to be an independent factor but during the course of the survey it behaved more and more as if determined by education. The explanation seemed to be that in recently married couples there was a delay of some months before they found work to match their educational achievements. The other traits were assorting independently. One might have expected, say, smoking and religion to show some convergence but no such tendency was evident in the overall sample. These social traits all play a part in selecting a partner therefore, but do not tend to converge or diverge markedly thereafter.

The main correlations found were in age, height, neuroticism, education, smoking habits, and religious practice. Assortative marriage for these traits may well play a role in cultural adaptation to change.

DECLARATION

I declare that the research described in this thesis is my own work. The work was carried out in the Department of Human Genetics, University of Edinburgh, and has not been submitted for any other degree.

G.C. Sutton
30th September 1980

A C K N O W L E D G E M E N T S

I am very grateful to Professor A.E.H. Emery for his constant encouragement and interest throughout the course of this study.

I also thank the Carnegie Trust for the Universities of Scotland for financial support throughout the study, and the Marie Stopes Fund of the Eugenics Society for assistance with research costs.

I am grateful to Frances Schofield for the typing of this script.

Finally, I am grateful to the citizens of Edinburgh who took part in this project.

CHAPTER 1.

Background

I. Introduction

II. Motivation: why study assortative marriage?

a) Social
b) Genetic
c) Practical criticisms

III. Measurements of r_{sp}

a) General comments
b) General traits
c) Physical traits
d) Psychological traits
e) Pathological traits

CHAPTER ONE

1. INTRODUCTION

Of the factors influencing the genetic structure of modern Western populations - mutation, drift, inbreeding, flow, selection, and assortative marriage - all but the last are of limited effect. The first three have little influence and the fourth and fifth mostly act slowly, over many generations. By contrast assortative marriage or assortative mating (AM) has potentially large effects on the population over a short period of time. Moreover it has social as well as purely genetic implications.

AM is defined as a tendency for married partners to have a greater similarity than would be expected by chance. For example, people usually choose partners of similar age and social background. In the present study, and in the previous studies referred to, it is assortative marriage, rather than mating, which will be discussed, as it is the marriage, not the offspring, which is the basis of ascertainment and focus of attention. However, when we discuss the implications for the genetic and phenotypic structure of the next generation we are implying assortative mating, rather than marriage. This is a rather fine point of distinction and the terms are generally used synonymously in the literature.

The <u>interspouse correlation coefficient</u> (r_{sp} or r) is the measure of the strength of AM for a particular characteristic. A value r = 1 indicates that partners are identical with respect to the trait. The value r = 0 indicates no similarity beyond that expected by chance. A negative value, to the limit r = -1, indicates disassortative mating, a tendency for partners to be dissimilar. An equivalent interpretation of r (which does not reflect the real world but is sometimes mathematically convenient) is that a proportion r of the population are matched assortatively, and that the remainder 1-r are paired off at random. The method of calculating r is given in standard statistical texts, e.g. Snedecor & Cochran (1967).

<u>AM versus inbreeding</u>. AM is a phenotypic process. It affects the genotype of the offspring and hence of the next generation of partners, but only if the characteristic is to some extent genetically determined. In contrast, inbreeding, which also induces similarities between partners, is primarily genotypic and only secondarily phenotypic. AM induces similarity for specific characteristics, whereas inbreeding affects the entire genotype.

II. MOTIVATION : WHY STUDY AM?

Information on AM is of both social and genetic interest.

a) Social: AM may be one of the mechanisms whereby social structures are stabilised and perpetuated.

- The degree of AM, especially for psychosocial traits, may be relevant to the success of the partnership.

- AM may greatly increase the number of people with extreme values of a trait. Thus, for instance, it is claimed that the number of people whose IQ exceeds 150 is roughly trebled by AM, and so is the number whose IQ is below 50. Both these sets of extremes will have an effect on society out of proportion to their numbers.

- AM for social traits may facilitate the passing on to the offspring of factors such as values, expectations, and attitudes, which are not genetic but are strongly familial.

- Knowledge of the strength of AM permits testing of hypotheses about why people attract each other. For detailed discussion see Eckland (1968).

b) Genetic: AM is a departure from random breeding. It therefore disturbs the Hardy-Weinburg equilibrium, and alters the genetic constitution of the population. The effect of this depends on whether the trait is Mendelian recessive, or multifactorial.

1) Mendelian recessive diseases. AM increases homozygosity at the locus concerned. Therefore the incidence of recessive diseases is increased (but see page 8). Specifically, if p and q are dominant and recessive gene frequency respectively, and I the incidence of the disease, then:

- with random mating, $I = q^2$

- with AM $$I = \frac{q^2}{1-rp^2}$$

Thus if q= 0.2 and r = 0.2, incidence would rise from 4% to 4.6%. The above formula is a shortened version of one given by Crow & Felsenstein (1968). It is an accurate approximation when q and r are less than 0.5, i.e. under all realistic conditions. In the hypothetical situation where r = 1, then:

$$I = q$$

Thus if q = 0.2, incidence would rise from 4% to 20%. (There are a number of practical criticisms to be made of these and the other formulae presented in this section. These criticisms will be discussed in section c) below).

2. Multifactorial traits. The effects of AM will depend on h^2, the heritability of the trait, and **n**, the number of genetic loci affecting the trait.

- AM increases homozygosity at the loci concerned. The effect is small when **n** is large and h^2 is low, which is probably the case for most of the traits reviewed here. The effect of AM is large when **n** is small and h^2 is high, as is the case for skin colour.

- AM increases the variance of the total population, because there are more representatives of the two homozygous extremes. This is particularly noticeable when **n** is large, and should be a significant effect in the traits to be discussed. The effect is cumulative over several generations leading to a new equilibrium frequency distribution more scattered than the standard Normal distribution. The increase in variance may be predicted by the following method:

Given r_{sp}, interspouse correlation coefficient
r_{po}, parent-offspring correlation
r_{ss}, sib-sib correlation among offspring

then the total variance of the trait may be partitioned into environmental and genetic portions, v_{env} and v_{gen}, which are assumed (unrealistically) not to interact. The genetic portion further divides into dominance and additive components, V_{dom} and V_{add}; and the additive portion subdivides into that due to random mating (V_{rand}) and that extra variance induced by assortative mating (V_{am}).

Then $\quad V_{add} = 2r_{po} / (1 + r_{sp})$

and $\quad A = V_{add} \times r_{sp}$

and $\quad V_{gen} = 4r_{ss} - V_{add} \times (1 + 2A)$

Hence calculate $\quad V_{env} = 1 - V_{gen}$

$$V_{dom} = V_{gen} - V_{add}$$

$$V_{rand} = V_{add} \times (1 - A)$$

$$V_{am} = V_{add} \times A$$

These formulae were derived from Burt and Howard (1956) on the basis of Burt's work on the IQ of London schoolchildren. That experimental work has now been discredited but the above theoretical approach is still accepted as valid.

The following worked example is given by Emery (1976): If in the case of intelligence it is found that $r_{sp} = 0.39$, $r_{po} = 0.49$, $r_{ss} = 0.51$, then it can be calculated that $V_{env} = 7\%$, $V_{dom} = 23\%$, $V_{add} = 51\%$, and $V_{am} = 19\%$. This 19% increase in the variance of the distribution of intelligence would treble the number of people whose IQ exceeds 150.

- AM increases the number of people with multifactorial diseases. These are assumed to be determined by an underlying continuously distributed liability, but to have a threshold for expression. If variance is increased, the proportion of the population which deviates from the mean by any arbitrary threshold amount, will also be increased.

If heritability $h^2 = 100\%$, then at the new equilibrium under assortative mating: $\quad I = r / [2n_e (1-r) + r]$ where n_e is the "equivalent number of genes", the number of equally frequent genes which would have the same effect as the actual number n, whose frequencies are unknown, but presumably unequal. (Crow & Felsenstein 1968).

- AM does not of itself change gene frequencies. However if directional selection is occurring, the response to it is faster. (This also applies to Mendelian recessive conditions.) AM may lead to relaxation of normalising selection, as it increases the likelihood that the more atypical members of the population will be able to find a suitable partner.

- In other species, AM may be an isolating mechanism relevant to species formation. For example fruit flies may separate into evening and morning breeders, and subsequently diverge further. However to cause such speciation AM must be very rigid and long-standing. Even a very small amount of gene interflow will prevent divergence. So while this process may be relevant in the stereotyped world of the insect, it is unlikely that it has ever been relevant to man, even in the most rigid human breeding systems such as the Indian caste system.

c). Practical criticisms of the above formulae and arguments.

There are many assumptions implicit in the multifactorial model. For instance it is assumed that a multifactorial trait is determined by the additive interaction of many small factors, so the central limit theorem applies and the trait shows a Normal distribution. There can be no non-additive interactioneither between genotype and environment, or between genes (such as epistasis).

In traits such as height, it is accepted that the multifactorial model is a useful tool for practical purposes even if its conditions are not rigidly met. However in mental traits such as intelligence and personality factors, which are in many ways more interesting, it is debatable if the model applies. The interaction of environmental and genetic factors is particularly complex here and is liable to confound attempts to measure heritability h^2.

There is a danger of circular reasoning. The formulae are presented in a way that implies that given various parameters such as n, h^2, q, and r, we can predict I the incidence of the condition. In practice the reverse is the case. We know I, r, and sometimes h^2, and use these to infer values of n and q. For example in the case of recessive deafness a value of roughly $n = 35$ has been deduced from study of offspring of deaf parents (Fraser 1976). Obviously we must be cautious in using that value to predict the future behaviour of deafness under varying conditions.

Man is a notoriously intractable subject for study. It is relatively straight-forward to measure r_{sp} itself. However the long generation time makes it very difficult to test predictions about the effect of AM.

In defence of the arguments set out in section b) above it may be said that the predicted effects on the offspring are intuitively reasonable. They have in fact been partly confirmed by experiment in other species, e.g. by McBride & Robertson (1963) in Drosophila. In this paper the emphasis will be on empirical results, rather than theoretical predictions.

I I. PREVIOUS MEASUREMENTS OF r_{sp}

a). General Comments.

Assortative marriage has previously been studied in a great variety of traits. Extensive catalogues have been drawn up by Spuhler (1968) and Roberts (1977) of physical and general traits, by Spuhler (1967) of psychological traits; and various psychosocial factors are reviewed by Vandenberg (1972), Wilson & Nias (1976) and Coleman (1977). Practically every measurable trait has been examined from colour of hair (9 studies) to size of foot (3 studies). What is one to make of correlations for ear lobe length, forehead skin reflectance, chest diameter or xipho-epigastric length, all of which have been found to be significant at the level $p < 0.01$? Does this data mean anything at all?

Firstly, if $p < 0.05$ is chosen as the threshold for significance, then 5% of traits studied will show significant similarity by chance alone. However this can not account for the vast majority of recorded similarities, many of which are significant at the 1% level.

Secondly, many of the traits are intercorrelated; the physical ones with height, the psychosocial ones with social class and intelligence. For instance, when sitting height is corrected for total height, no significant degree of AM remains. A number of partial correlations have been undertaken, but many traits continue to show significant similarity even after these corrections have been made (e.g. forearm length, bicristal diameter, and arm length). Perhaps correction is being made for the wrong trait. Multivariate analysis is needed to clarify this point.

Thirdly, some of the techniques are not very repeatable. This inaccuracy is not fully reflected in the size of the standard deviation if only single measurements are made. For instance, bicondylar breadth at the elbow ($r = 0.28 \pm 0.064$) has poor repeatability. Its self-consistency coefficient (i.e. the correlation between two readings at the same site in the same person) is only 0.45 (R.G. Burwell, personal communication 1977). However bicondylar breadth at the knee has a much higher repeatability (0.95), and is significantly correlated with $r = 0.2 \pm 0.067$.

Fourthly, the nature of the sample, and the method of ascertainment, can critically affect the results. For instance Wilson and Nias (1976) describe one study apparently showing that whilst there is AM for physical attractiveness, there is none for intelligence. But the participants were all from an American university college: the variance due to intelligence would have been practically eliminated in the sample under study. Pomerat (1936) found unusually high similarity for height (0.63) and weight (0.30), comparitively low correlation for age (0.51). His sample was of 107 infertile couples, mostly with endocrine disease. In general the degree of similarity between couples is not related to their fertility (Spuhler 1968). Spuhler's study in Ann Arbor, Michigan, is atypical in finding strong correlation for various measures of head size. Presumably this sample, unlike most others, was racially mixed, in which case similarities would be secondary to AM for racial origin. In practically any experiment there will be a social class bias in those who are willing to take part, and this could also affect results.

Fifthly, it is possible that some similarities (eg. for psychological traits) may have arisen since the time of marriage. This could happen in more than one way. The couples may actually have become more similar the longer they lived together; or dissimilar couples may have tended to split up and be selectively lost to investigation, or the trait may have been correlated with age or some secular trend. Several investigations have attempted to relate degree of correlation to length of marriage, e.g. Johnson et al (1965) for blood pressure and serum cholesterol, Kreitman (1964) for personality factors, and Willoughby (1933) for weight. These studies have been unrevealing and in any case do not distinguish between the various mechanisms mentioned. An experimental design to overcome these problems would involve:

1) Careful attention to ascertainment. Inevitably there will be a social class bias in response rate and one must consider how this confounding effect might be overcome.
2) Choosing traits for study which are fairly easily measured, and for which there are intuitive reasons for expecting AM. Analysis of results will be much easier if parametric methods can be used: implying a reasonable sample size, and a suitable choice of scale for measurement.
3) Appropriate statistical techniques to determine which correlations are secondary to others.
4) Remeasurement of the sample after an interval of time, to see if each couple has converged, diverged or separated altogether.

5) Comparison of different marriage cohorts.

In discussing traits in more detail, it is useful to categorise them as general, physical, psychosocial and pathological.

b) General Traits

1. Age at Marriage: This usually shows very high correlations, with r in the region 0.8 to 0.9. The correlation coefficient ranged from 0.51 to 0.99 in 22 studies quoted by Roberts (1977). The lowest value was that of Pomerat's infertile couples. The highest value was from Spuhler's study of the Ramah Navaho, an Indian tribe which did not otherwise practise AM for physical traits. The high correlation for age means that a trait which altered with age or showed a secular trend could appear to show AM, secondary to the effect of age. Fertility (see below) is an example.

2. Fertility: The number of sibs of married partners has been shown to display significant AM. However there have been secular trends in fertility. Beckmann and Elston (1962) demonstrated that when the sample was controlled for date of birth, no significant correlation remained. Nowadays people can control their own fertility and it would be interesting to know how closely partners' plans are correlated.

c) Physical Traits

1) Height has generally shown significant correlations in the range 0.1 <r <0.3. This results in a marked increase in variance of the offspring. (V_{add-AM}= 17%, if analysis by the method of Burt is accepted). Values outside this range can mostly be explained by the nature of the sample. Small inbred societies, like the Ramah Navaho and the Solomon Islanders, have often shown non-significant values (summarised by Roberts 1977). In such societies, there may be very few available partners at any moment in time, and once those of the wrong age have been rejected there is no scope for choice regarding other traits. Johnston (1970) found a high correlation for height in the Cashinahua Indians of Peru (r = 0.346 $\pm$ 0.136) but here marriage occurred before

growth was complete. The correlation was therefore largely due to AM for age. The exceptional result of Pomerat ($r = 0.63 \pm 0.058$) was mentioned earlier. Why should there be AM for height? Social activities may be easier if partners are of roughly the same height, and social prejudice demands that the male be no shorter than the female. These are weak reasons; but AM for height is weak. More important perhaps than actual height is perceived height. People can alter their apparent height with platform heels and it would be interesting to know how this apparent height is related to the real or relative heights of the partners.

2) Weight: In Western populations, r_{sp} is generally significant, averaging about 0.2. The lowest value, from Ann Arbor, Michigan, ($r = 0.08 \pm 0.07$) was nonsignificant but curiously in the same sample weight at marriage was significant ($r = 0.23 \pm 0.066$). Perhaps weight diverges after marriage - maybe because the male puts on more weight in his thirties. In non-European populations the correlations were nonsignificant. Again the Cashinahua were atypical ($r = 0.313 \pm 0.14$).

3) Skinfold thickness has not been studied in Westerners. In non-Western populations it has shown no AM. Skinfold thickness is known to be correlated with total body fat (Durnin and Womersly 1974) and hence with weight. It is possible therefore that it would show AM in Western populations since they, unlike the societies studied to date, assort for weight. It could even be hypothesized that skinfold thickness, not weight, is the more readily perceived phenotypic trait, determining as it does the general roundness or angularity of the body.

4) Head size: Roberts (1977) catalogued 41 measurements of the size of the head and face. The bulk showed no AM. The Ann Arbor study was again atypical, finding a number of significant correlations, such as for interpalpebral breadth, inter-pupillary breadth, and ear length. Perhaps these correlations were secondary to AM for racial origin.

5) Physical attractiveness was shown, in a series of studies reviewed by Wilson & Nias (1976) to be a fairly objective trait (i.e. a series of observers would rank a group of subjects in much the same order) and to show AM. This conclusion is intuitively reasonable but note that most work was done by computer-dating experiments on American campuses. Variance due to IQ and social class would have been largely removed from the sample. AM for facial appearance was reported by Griffiths & Kunz (1973) but the experiment was a curious one: several hundred observers tried to match ten photographs into five couples. They did so more accurately than could have been expected by chance, but clearly the age of each partner must have influenced the results.

6) Pigmentation: Eye and hair colour have seldom shown AM. Results for skin pigmentation depend acutely on the nature of the population. The black-white dichotomy in, say, a US city carries many social connotations, and AM is to be expected. In fact Adams (1969) concluded that the previous trend towards miscegenation in the US had reversed, and the whites and blacks were repolarising. This is because in the case of skin colour, n is low, r is high, and h^2 is high. Therefore heterozygosity falls quickly. If there is indeed a trend against miscegenation it is likely to be very temporary.

Gradations of skin colour within a racially homogenous population may show AM. It has only been studied in non-Europeans, principally Sikhs and Solomon Islanders. Correlations varied with technique and the site of the skin studied, being most consistently significant for the forehead of the Sikh. These shades of colour may, perhaps, have social significance, and the correlations may have been secondary to AM for social class.

7) Physiological measurements: Blood pressure showed no AM. Pulse was significantly correlated before exercise ($r = 0.29 \pm 0.092$), but negatively afterwards ($r = -0.20 \pm 0.096$) for which there seems no obvious reason. Since pulse and BP are related to age, smoking, and physical fitness, one might have expected a secondary correlation at least.

8) <u>Cryptic physical traits</u>: A great number of other physical traits have been examined. Generally they have shown significant values of r; reasons for this have been discussed earlier. Dermatoglyphics and serum protein fractions were not correlated. If they had been, one would suspect inbreeding as the cause (and presumably it was to eliminate that possibility that investigators were studying them). One remarkable exception was IgA in Northumberland ($r = 0.703 \pm 0.092$). Perhaps some local viral infection was affecting results.

d) <u>Psychosocial Traits</u>.

1) <u>Intelligence</u>: Spuhler (1967) cites 7 studies between 1928 and 1946. The unweighted mean of r was 0.4, ranging from 0.03 ± 0.06 for "Arithmetic" to 0.76 for "Progressive Matrices". Different components of total IQ may vary in importance in different cultural groups.

Harrison et al (1976) found AM for total, verbal, and performance IQ, but these (and other traits studied) varied within the sample. There had been considerable recent migration into the area (Otmoor villages, Oxfordshire) and if both or neither of the partners were local, there was a correlation for IQ. If only one was local, no AM was found. No AM for IQ existed in social classes IV and V. By contrast no AM for height occurred in classes I, II and III. Overall, total IQ had $r = 0.279$. Note that in a rural sample, classes 4 and 5 would be mainly farm labourers: probably a fairly stable population. In a city, very different results might be expected. In fact Mascie-Taylor and Gibson (1979) found assortative marriage for IQ at all social levels in a Cambridge suburb.

Johnston et al (1976) found small correlations for specific cognitive abilities, but when the sample was controlled for educational attainment the similarities were nonsignificant. Any attempt in such studies to control IQ for education leads to a dilemma. High IQ leads to higher education, but education leads to practice at the sort of abilities involved in the IQ test, which will increase the IQ score.

One rather exotic result was that of Guttman (1970) who studied "subitizing" - the ability to tell at a glance how many objects are present in a small group, h^2 for this trait was 0.60, r_{sp} was between 0.46 and 0.62. The trait was said to be distinct from "estimating" where the group is large: estimating is not heritable. The population studied were Israeli Sephardim, whose cousin marriage rate was 5 to 10%. So this correlation may reflect inbreeding, or be culture specific. Confining our attention to those traits composing the standard IQ tests, Spuhler's own results from Ann Arbor were:

"Progressive Matrices" r = 0.399

"Chicago Verbal" : total right r = 0.305

"Chicago Verbal" : percentage right r = 0.732.

2) Personality traits: This is one of the few areas where one might expect to find disassortative mating, as for some traits (eg dominance/submissiveness) people might choose a complementary partner. In fact, complementation does not appear to occur. If people display any particular preference with regard to a personality trait, it is towards similar partners.

Spuhler (1967) cited 15 values from 5 studies between 1909 and 1946. The unweighted mean was r = 0.14, range 0.02 - 0.29, but this included some curious traits like "criminality" and "insanity" which are not part of modern standard tests. Wilson & Nias (1976) concluded that personality factors did not, in general, show AM. Harrison et al (1976) found no AM for neuroticism or extroversion, but r = 0.159 for "personal inconsistency". Again the picture here was confused by immigration. The correlations were higher in those married since the war. The most recently married cohort consisted of those wed from 1960 onwards. Hence it is not possible to see what is happening in the first five to ten years of marriage.

It would be very difficult to ascertain and study a large sample of divorced couples, but it would be fascinating to know if couples who split up had different initial correlations.

3) Social Class: Studies are seldom strictly comparable because class may be defined in so many ways. Occupation is the obvious method but a woman's occupation at the time she marries is often a poor guide to her status. Income, education, area of residence, and father's occupation are other measures, which are highly intercorrelated. Investigators tend not to quote values of r but a different measure called "C", which is not so easily interpreted, or χ^2. There is a case for saying that in a reasonable sized sample, the method of calculating r is sufficiently robust that it does not matter that "class" is a non-parametric trait. Certainly all these methods demonstrate that the degree of AM is too large to be accounted for by chance. Thus Kiser (1968) showed clear evidence of AM for social class, but Warren (1966) found that little AM for occupation remained after he had corrected for education. The only large study of unmarried couples showed high correlation for religion, family background, and ideas about ideal family size. The sample was 1000 middle-class engaged couples in pre-war Chicago (Burgess & Wallin 1944). More recently in Reading, Coleman (1977) confirmed that the strongest correlations were for religious affiliations and practice.

There appears to be a secular trend here: couples married in recent years show a lower correlation for religion than those married earlier this century.

Also in Reading, education showed correlations of around r = 0.58 for upper social classes, r = 0.25 for middle and lower classes. The correlation may in part have been secondary to the effect of age, as educational opportunities have increased greatly since the war.

Some attempt has been made to study distance between residence of partners before marriage. These are difficult to interpret because of the idiosyncracies of human dwelling patterns. Social distance does not equal bee-line distance, even in a rectilinear American city. It must also be remembered that in some societies a substantial percentage of couples have been living together for some time before marrying.

e) Pathological Traits

Age at death showed a correlation of about r = 0.3 but secular trends in life expectancy may account for most of this. There are a number of disease processes for which AM may be operating. If the disease is partly genetic, and if there is appreciable morbidity and mortality, then selection may occur. Therefore, the relevant genes will be removed from the population more rapidly than would occur under random mating: a long term eugenic effect. However the number of cases will meanwhile be raised : a short term dysgenic effect. The short-term disadvantage is of greater practical importance than the long-term advantage.

1) Speculative examples: It might be predicted that organisations to promote the interests of, say, epileptics, and to hold social gatherings for them, might lead to an increase in the number of marriages between epileptics. Many causes of epilepsy are non-genetic, but a sizeable proportion of cases of "idiopathic" childhood epilepsy are familial and may well be genetic. Better treatment means that they are more likely to eventually marry and have children themselves. In that case a marked increase in disease incidence could be predicted.

Similar considerations might apply to ischaemic heart disease which is now affecting younger people. The effect might be compounded by AM for social class, implying similar smoking habits and tendency to eat an atherogenic diet. The effect of AM, however, would probably be over-shadowed by that of improved prognosis, screening and preventative programs.

2) Real examples: AM is found in deafness (Fraser 1976). Deaf partners can communicate better with each other than with normal people, so a concordantly deaf marriage will be more successful than a deaf with normal. However, because so many genes and environmental phenocopies are involved there is little effect on population frequency. (It was pointed out earlier that it is the number of genes that is inferred from trait frequency, not vice versa). AM is only significant if the genes are

Mendelian recessive or additive. In Britain many syndromes are in fact recessive, and inbreeding also contributes to their frequency. Fraser (1976) claims that by contrast, in South Australia syndromes are often Mendelian dominant, so neither AM nor inbreeding are relevant there.

AM for IQ occurs, and if the example given earlier is to be believed then this trebles the number of people whose IQ is below 50. Hence AM could contribute significantly to the prevalence of mental handicap. However most cases of IQ $<$ 50 result from simple pathological processes, and the contribution from the tail of the Normal distribution is small. Also the lowest IQ people usually fail to marry and are therefore subfertile.

Psychiatric disturbance occurs in both spouses more often than expected by chance; especially if it is a case of neurosis rather than psychosis. This could be AM, but it may also be through behaviour copying by the previously well partner, or a nonspecific reaction to the stress of having a mentally ill spouse. It could be even an artefact, with the referral of one partner bringing the other to medical attention.

Ovenstone (1973) showed that a nonspecific reaction was the likeliest explanation. There was no similarity of pre-morbid personality, and partners were not, on the whole, similar for specific symptoms. Kreitman (1964) showed that a neurotic group had little correlation for mental health when they married, whereas normal people had high correlations. Thereafter normal values fell somewhat with time, whilst patient values rose.

Thus AM for disease processes is not currently an important factor in their incidence. It might conceivably become important in future, if removal of environmental causes meant that more cases were genetic. Then h^2 would become high, and improved treatment would mean that patients remained fertile. The strength of AM would then influence the frequency of the disease.

CHAPTER 2.

The Present Study: Aims & Methods.

I. Aims of the Study

II. Methodology

a) Experimental design

b) Ascertainment

c) Recruitment to the study.

d) The first visit

e) Encoding and extra visits

f) The second visit

g) Sources of bias

h) The computer program DANTE

CHAPTER TWO

I. AIMS OF THE STUDY

a) To establish the strength of assortative marriage, by measuring the interspouse correlation coefficient r_{sp} for a number of physical, psychological, social and general characteristics, in a series of Edinburgh couples.

b) To discover to what extent any similarities found were present at the time of marriage, or whether the couple have converged through shared domestic environment.

c) To unravel, by appropriate mathematical techniques, the interaction of those traits which are strongly associated and which thus give rise to spurious correlations.

d) To assess the implications of the similarities found for those genetic and social processes which assortative marriage is believed to influence.

I I. METHODOLOGY

a) Experimental Design

The question arose whether to study a few traits in many couples, or many traits in a few couples. If few traits are studied then one can not discover whether they are interacting with the other (unmeasured) traits. But if few couples are studied then the results are excessively susceptible to the vagaries of sampling. An attempt has been made in this study to gain the best of both worlds, by including a wide selection of traits, all of which can be swiftly measured, thus allowing a large sample of couples to be studied.

Although the original motivation for this study was genetic, it was the social implications of assortative marriage that seemed of greatest interest. Moreover, those traits which were known to be largely genetic (such as height) had been extensively examined in the past. By contrast less was known about those social and general traits which were familial but non-genetic. For instance, no previous study of smoking habits was found. An attempt was therefore made to cover as many social, psychological and general traits as possible, as well as those anthropometric variables which might have social consequences. A completely different experimental design would have been necessary to investigate pathological traits so these were not included.

If couples are converging or diverging, one would expect most changes to occur at two different times: shortly after marriage, and at the birth of the first child. These are the beginnings of the first and second stages of marriage defined by Dominian (1979). Therefore two cohorts were chosen - a group who had been married less than a year, and a group who had been married five or more years. It was intended to study all these couples twice, with a year's time lapse. It was possible that, in the case of the newly-weds, some of the similarities observed could have arisen even in the short time they had been together. Therefore a third cohort was chosen, of couples who had not yet married, and these were restudied a short time after marriage.

Since a wide selection of traits were being measured on a large sample of couples, a considerable body of data was amassed, and the aid of a computer had to be enlisted. This was necessary not just for calculating correlations, but for displaying the data in a comprehensible manner and for testing various hypotheses that suggested themselves from time to time. Although a number of general purpose statistical computer programs have been published, they appeared rather inflexible. Therefore an interactive program was designed for the task at hand. It is described in a later section of this chapter.

b) Ascertainment

Edinburgh: The city boundaries have recently been altered. For the purpose of this study, "Edinburgh" should be understood to mean "The City and County of Edinburgh" as defined by the Edinburgh Corporation Act 1920. This includes the conurbation that most citizens would think of as Edinburgh: the city itself and Leith. It consists of seven registry districts of Newington, Morningside, Haymarket, George Square (formerly St. Giles), St. Andrews, Canongate/Portobello, and Leith. It excludes Musselburgh, Dalkeith, Bonnyrigg, Lasswade, Currie, Kirkliston and South Queensferry.

Newly Weds: Couples were ascertained from registry office notices, and from wedding announcements in the newspapers.

Until the marriage laws of Scotland were changed, at the end of 1977, couples wishing to marry had to give notice of their intention to do so in one of two ways: by notice displayed at the registrar's office of the district in which they lived, or by banns read out in church. About 10% of couples gave banns in a church. This was only in the case of

Church of Scotland marriages. Members of other denominations had to give notice at a registrar's office even if they were to marry in church. Church banns were read out at services and no written notice posted. Therefore these couples could only be ascertained via the local newspapers. Couples' names and addresses were collected from the Scotsman and Evening News between 19th October 1977 and 14th January 1978. 57 pairs were found.

The other 90% gave notice via the registrar's office. The notices were displayed for eight days and gave the full name, address, occupation and previous marital status of each partner. Each registrar's office was therefore visited once a week and the information recorded. This meant that the social-class profile and previous marital status of almost the entire marrying population was known, and this could be compared with the profile of those who agreed to take part in the study.

Couples were collected between 19th October 1977 and 31st December 1977 and 498 pairs were found. Of these, 17 pairs were also found via the newspapers.

538 pairs of newly-weds were ascertained: 481 from registrars' offices alone, 40 from newspapers alone, and 17 from both sources. On 1st January 1978 the marriage laws of Scotland changed. All couples thereafter had to give notice via a registrar's office, but the notices simply gave the names and did not give address or occupation., although that information would still be available from the marriage certificate. The end of the year was therefore a logical point to stop but it seemed likely that church marriages had been incompletely ascertained. Therefore newspaper announcements were still collected for another two weeks; Hogmanay produced a small spate of announcements which then dried up as the month went on.

Married 5 years or more: All marriages in Scotland are eventually filed in the Record Office at Register House, Edinburgh. They are filed into volumes according to year. Within each year they are listed alphabetically by name of husband; cross-volumes also list them by name of wife. Each entry contains, against the husband's name, the wife's name, the registration district in which they were married, and the entry number of the marriage in the register for that district and year. From this entry number it is possible to find the record of the original marriage certificate giving occupation and address of both partners. This is the method used in tracing ancestry. However it was not practicable for this study because of the large numbers of people to be traced (the list of marriages for 1972 occupies 4 volumes, each about 2 inches thick) and because the couple's address would almost certainly have changed. The following search procedure was adopted instead. On finding a marriage taking place in one of the seven Edinburgh registry districts, the husband's name was looked for in the 1977 telephone directory. If one (and only one) person of that name was listed, and if he had an Edinburgh address, that directory entry was ticked as a "find". Of 4257 marriages in Edinburgh that year, 783 pairs were "found".

Engaged: Names were collected from the engagement columns of the Evening News and Scotsman, between 19th October 1977 and 1st May 1979. 350 pairs of names were recorded where an Edinburgh address was given for one or both partners.

A number of sources of bias are evident at this stage. They will be discussed in a later section.

c) Recruitment to the study

The 1671 couples whose names had been collected were each sent a letter, briefly explaining the nature of the study and inviting them to take part. A postcard was enclosed, suggesting a time and date when they might care to be visited at home. This was to be returned stating if the time was convenient. The letter and postcard are reproduced in the appendix. The following points may be noted:

1) The letter was Xeroxed not stencilled. This was more costly but gave a better quality letter which might improve response rates. A flexowriter would have been ideal but was not available.
2) The name of the couple was handwritten into each individual letter, rather than beginning with a drab "Dear Sir". Letters were also signed and dated individually.
3) The letter sent to the 1972 couples differed slightly from that sent to the newly weds and engaged: it made no reference to couples. This is because in the former case, the addressee might not have been married - being divorced or widowed or identified by mistake from the register - and might have been upset at receiving an inappropriate circular.
4) The postcard for the reply was pre-paid.
5) Weekday evenings were suggested, to catch both partners in.

It was decided to contact all the newly-weds first, so that they could be seen as soon as possible after marriage. These letters went out between February and June 1978. Thereafter letters were sent to the 1972 couples, so that most had been married for six or seven years by the time they were contacted (July 1978 to March 1979).

Engaged couples were contacted as soon as practicable after the engagement notice appeared, i.e. between February 1978 and May 1979.

A number of couples agreed to take part but had already moved out of town; these were excluded. Also excluded were those couples where it was only possible to visit one partner, or where they were already separated.

In the case of the 1972 couples, it was obvious that there would be several cases of mistaken identity. This raised a dilemma. If the "mistakes" were excluded then the survey might be short of couples. On the other hand, if they were included they might make the results unrepresentative - for instance the "mistakes" could be much older.

Initially it seemed that there would be a problem with numbers, so the "mistakes" were included. But after working with 1972 couples for some weeks it became evident that there would be no shortage of numbers. The response rate for this group was much higher than for the others, which had not been anticipated. Indeed it seemed that numbers would be so great in this group that the project would overrun the available time. Therefore after September 1978 no more "mistakes" were included. Those "mistakes" already visited remained in the study; account will be taken of their effect in presenting the final results.

A new edition of the Edinburgh phone directory was published in June 1979. Several couples already written to, but whose letter was returned by the GPO marked 'Gone Away', were traced from the new listing and recontacted.

Thus by the time recruitment to the study was complete, 81 engaged couples, 98 newly-weds, and 224 older couples had agreed to take part and had been visited at home.

d) The First Visit.

The couples were thanked for agreeing to take part. The purpose of the study was explained and any questions were answered (the two commonest being "How did you get our names?" and "What do you hope to gain from this information?").

The first part of the interview involved questions on social and general factors. Each partner was asked his or her name (plus original surname in the case of women) and date of birth. They were then asked their occupations. Occupations were recorded in enough detail to allow classification by the Registrar General's classification of social classes. Also recorded were the occupation at the time of marriage (or when engaged, in the case of the unmarried couples) and father's occupation.

Next they were asked at what age they had left school, and whether they had any O or H levels. Those with Highers were asked about degrees. General qualifications acquired since leaving school were taken into account (eg. O levels, general Open University degrees) but not specialist qualifications or those which could only be regarded as in-job training (MICE, HND, City & Guilds, nursing or secretarial studies). Nor were diplomas included as it was hard to tell what standard they imply.

Engaged couples were asked when they had become engaged and when they planned to marry. The others were asked when they had married, whether they had been engaged, and if so whether the engagement had been announced in the newspapers. This was to try to discover how typical the engaged couples, who could only be ascertained via the newspapers, were of all the couples.

Partners were asked if they had ever been married before. If so they were asked if they were divorced or widowed. Also any remarriages or separations by their parents were recorded.

They were asked how many brothers and sisters they themselves had, as well as brothers and sisters of their father and mother (not aunts and uncles who are not necessarily true 'relatives'). Adopted and foster sibs were counted but half sibs were not. Sibs who died before the age of one, and parents' sibs who died before age ten, were not counted. When a parent had remarried, partners were asked which set of parents and step-parents they had mostly grown up with and regarded as 'their own', and the questions about sibs rephrased to apply to that pair. (Occasionally at this stage a step-father's occupation might be substituted for that of the natural father in that section).

Any partner who had been married before was asked if he/she had children by that marriage. The 1972 couples, and some newly-weds, were asked if they had children by the present marriage. Unless the woman was aged 40 or more, partners were asked: "The answer to the next question does not commit you to anything. Do you have any views on how many of a family you would want altogether? . . . . . do you have any views on when you would want a family - not necessarily in terms of time, but just in terms of what's going to decide you; money, house, job, etcetera?" (The phrasing is important. Several variants were tried but none elicited the information quite as well.) The man was asked these questions first as it seemed likely that the woman's opinion would be decisive. No questions were asked about birth control, although it was occasionally volunteered that, say, one partner had been sterilised.

Partners were then asked how much they smoked, then about religion (again the phrasing is important): "Do you currently practise any particular religion? The definition of 'practise' is up to you, it's purely your own opinion of yourself that counts."

The second part of the interview involved physical measurements. By this stage the couple had been sitting quietly for several minutes.

Resting pulse was measured over 30 seconds.

Blood pressure was recorded, with the person sitting. Precise details of technique of this and the other physical measurements are given in the appropriate sections of Chapter 4.

Couples were then asked, "How tall do you think you are?" and "How much do you think you weigh?" Women who were pregnant or recently delivered were asked how much they had weighed before pregnancy. After this information had been recorded, the actual height and weight were also measured.

An exercise tolerance test was then performed. It involved standing up and sitting down, on a standard firm chair, ten times, while carrying a weight in the outstretched arms. After ten seconds rest, the pulse was recorded over fifteen seconds.

Skinfold thickness was measured over the left triceps, left biceps, chin, and behind the left knee.

The third part of the interview involved a personality test.

Partners had to assess themselves on the Eysenck Personality Inventory (see Chap. 5). Having done so, they were told "Now answer the same questions again as you think they apply to your partner. Not what you think your partner would answer, but what you think is true about your partner."

Most newly-weds, and some of the 1972 couples, were asked if they had been living together before they were married. This was not asked if it seemed likely to be resented. If they had lived together for some time, they were asked if their occupations at the time they moved in together were the same as when they finally married. In the case of the 1972 couples so few had been cohabiting - and those few actually volunteered the information without being asked - that this question was dropped.

At the outset of the study, the inclusion of a fourth group - separated couples - was considered. Several partners who had been divorced were asked if they would object to their ex-spouse being contacted. In general people were surprisingly willing to allow this to be done. However it was obvious that tracing would be difficult, response rate probably low, and it was unlikely that an adequate sample could be collected. The idea was therefore abandoned and no attempt made to contact ex-spouses.

Partners were then asked if there had been any change in their family surname - either in their own or their father's time. Each couple was also asked if they knew of any common relatives such as distant cousins by which they might be related.

Finally, couples were thanked for their help, and asked if they were willing to be visited again in a year's time. If there was a possibility that they would have moved house by then, a forwarding address was noted.

e) Encoding and extra visits

Data were encoded the following morning and eventually filed on the computer. But it was evident that some couples would have to be classified into cohorts different from that in which they had been originally ascertained. Some engaged couples married before they could be visited, and were treated as newly-weds. Some couples cohabited before they married, causing engaged couples to be moved into the newly-wed category, and newly-weds to be treated as 1972 marriages. Some older couples had split up and the man had recently remarried. Some couples were ascertained by mistake. The numbers in each category are detailed in Figure A, which shows the three stages of ascertainment, agreeing to take part and reallocation to different groups. Each couple was assigned to one of the four following cohorts:-

i) ENGAGED - not yet married: 68 pairs

ii) NEW - living together less than $4\frac{1}{2}$ years when visited: 113 pairs.

iii) OLD - living together $4\frac{1}{2}$ to $9\frac{1}{2}$ years when visited: 205 pairs.

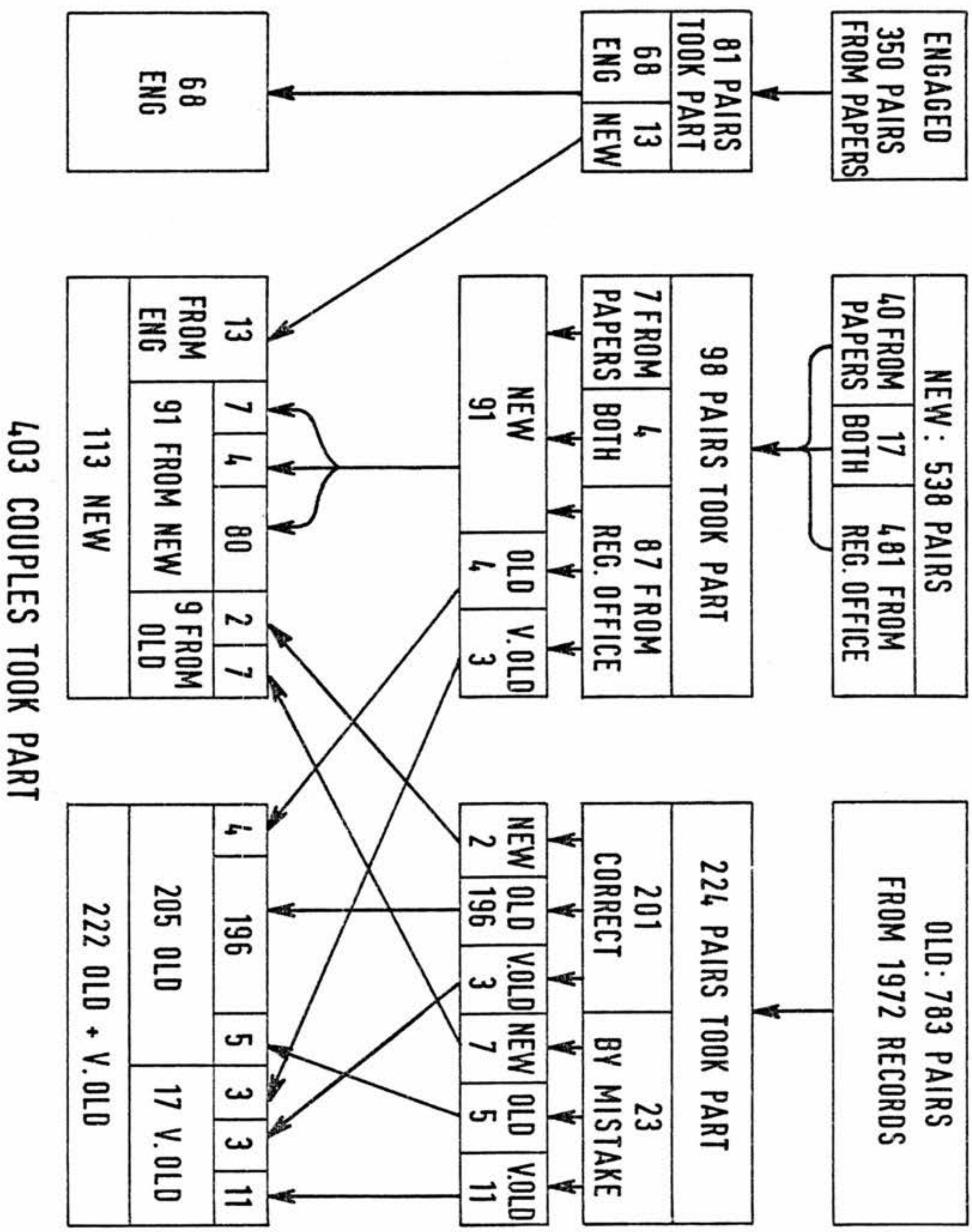
iv) VERY OLD - living together ten years or more: 17 pairs. Since this is a very small group the 'very old' have been lumped with the 'old' for the rest of the analysis.

The composition of each group is shown in Figure A.

Of course the terms 'old' and 'very old' did not refer to the ages of the partners but to the length of the marriage. Even so they were only old relative to the newly-wed category but the first decade or so of marriage was of particular interest in this study.

There was a delay in receiving the skin-fold calipers from the manufacturers. By the time they arrived some 30 couples had already been visited. These were all recontacted (see letter in Appendix) and as many as possible briefly visited to take the measurements. This also allowed correction of any ambiguities or omissions in the data which inevitably occurred in the first few couples before the interview technique became polished.

FIGURE A
ASCERTAINMENT & RECRUITMENT



f) The Second Visit

About one year after the first visit, new and old couples were recontacted. Engaged couples were recontacted about two months after the planned wedding date. The letter sent is in the appendix.

At the second visit, it was first checked that dates of birth had been correctly recorded and whether occupations had changed. Previously 'Engaged' couples were asked the date of marriage. Smoking habits and current religious practice were then recorded. Because the definition of active religious practice was left to the individual, answers might vary although habits had not altered. Therefore if the answer differed from the previous year, the person was asked if they thought there had been a change over the year, or whether it was simply the way the question had been asked that had made them answer differently. Thus, at this point the person's religious practice for the previous year sometimes had to be reclassified. Unless it was obvious that the couple would have had no more children (because both had expressed a strong preference for no more on the previous visit) then they were asked if they had had any more in the interim, and how many more were contemplated and when. Again, the man was asked before the woman. Then followed, as before, pulse and blood pressure at rest, how heavy they thought they were, height and weight, pulse after exercise, skinfold thickness, and the personality test (assessment of self and of partner).

Some engaged couples were then asked whether they had been living together before they had married, and if so when they had moved in together. Basically this was asked of those who had their own flats and were not living with parents at the time of the first visit. Partners were also asked if they were left or right-handed.

Finally, they were thanked for taking part in the survey.

Some of the engaged couples had to be retrospectively reclassified as 'new' because they were already living together at the time of the first visit. Those couples have been taken into account in the numbers in Table 2.1.

Couples who had moved out of town in the intervening year were restudied if this was logistically possible. Often, for example, they would come to Edinburgh to revisit parents. Inevitably however some were lost to follow-up by moving abroad. A few couples did not wish to be revisited.

Couples were not revisited when it was learned that there had been a death or separation. However these were not regarded as being 'lost to follow-up' as all marriages must end in one way or the other. The outcome of all the revisits is summarised in Table 2.1.

g) Sources of Bias

A number of sources of bias may be identified at this stage.

1. Couples who moved out of Edinburgh before they could be contacted were lost to study, and those who married elsewhere and later moved into town were never ascertained. Therefore there is a bias against those in highly mobile occupations. Couples were lost through moving between the first and second visits, but that was less of a problem as it was known exactly what sort of people had been lost.
2. Couples married in 1972 had to be listed in the Edinburgh telephone directory in 1977 in order to be 'found'. Those with no phone were lost, which may have caused a social class bias. This bias also operated to some extent in the other groups as telephone directories were frequently used to track down those who had moved.
3. Men married in 1972 had to have a reasonably distinctive name. Telephone directories seldom list Christian names, so those people with single initials were unlikely to be unambiguously identified.

TABLE 2.1

REVISITS UP TO 31st AUGUST 1980

	ENG	NEW	OLD
TOTAL	68	113	222
DUE TO HAVE BEEN REVISITED	61	113	221
TEMPORARILY LOST (no reply, gone away, etc.)	3	1	5
PERMANENTLY LOST (abroad, untraceable)	1	6	2
SPLIT UP	2	4	2
DIED	1	0	3
REFUSED	2	3	10
REVISITED	52	99	197

The more common the pattern of initials, the more unusual the surname had to be, and vice-versa. It was expected that this would cause a bias against the native Scots/English population, in favour of people with exotic surnames. In practice there seemed to be little evidence of this, with the correct "Fraser", "McGregor" and "Brown" (out of so many) being found and visited, and the wrong "Neri", "Szponar" and "Jagla" being excluded.

4. There was incomplete ascertainment of those giving banns in church, implying a bias against those who were practising members of the Church of Scotland. Probably these missed marriages were first marriages, and predominantly middle class.

5. The newly wed couples were all recruited in the last quarter of 1977. It is possible that this time-slice was atypical in several ways. Firstly, it missed the traditional surge of summer weddings. In some ways this may have been fortunate as in Edinburgh there would always be a large number of weddings of recently graduated students, which would have biased the sample. Secondly, the divorce laws had recently been changed, so that living apart for five years was a sufficient condition for divorce whether the other partner objected or not (or even knew about it). This may well have led to an excess of second marriages by people whose divorce had finally been granted. Thirdly, Building Society interest rates were unusually low during that period, so that many young couples who could not previously afford it might have been able to buy a house. Perhaps there was an excess of lower-middle class marriages on that account.

6. Those who had not yet married could only be ascertained if they became formally engaged and announced the fact in a newspaper, possibly an atypical thing to do. (Note however that The Scotsman's "Forthcoming Marriages" column was not used.) Some couples had planned very long engagements and were still unmarried at the end of this study.

7. Couples had to agree to take part. Biases in response rate must completely over-shadow every other factor. This bias will have been in social class, and perhaps in personality traits. There may also have been a bias towards people who are particularly interested in medical research - either as part of their job, or because of an illness. To a large extent it was possible to control for the social class bias by statistical analysis, and by comparison with the social class profile of all the newly weds ascertained. (But see 5 above).
8. Biases arose during the visits through instrument and observer error, e.g. digit-preference in recording blood pressure. These are discussed in the appropriate chapters later.
9. Edinburgh itself is an atypical town. It is a seat of government and university centre rather than a site of heavy industry. Therefore it has attracted few immigrants from overseas, but has many students. This is not exactly a 'bias' but a difference between Edinburgh and other British cities which must be taken into account before generalising any conclusions drawn from a local sample.

h) The Computer Program DANTE

There were 403 couples and, counting the 'his' and 'hers' values separately, some 150 traits to be analysed. A computer was clearly essential, and the program DANTE (see appendix) was written to handle the data.

The computer language used was IMP (ERCC 1970). The program was implemented on the ICL 2980 Edinburgh Multi-Access System (EMAS) and occupied 53K of store. In addition, the data files occupied 97K (old couples), 50K (new couples) and 30K (engaged couples). The statistical calculations performed by this program were actually of a fairly simple nature. It was the ability of the program to display the data in various ways, and to allow rapid hypothesis-testing, which was its main advantage.

The program is entered by typing RUN (DANTE) and in reply to INPUT: the user types the group of couples to be analysed (NEW, OLD or ENG). After a pause to read in this data from store, the program is ready for commands. The commands available in the current version of the program (MK 10) are as follows:

1. CORREL. To calculate the correlation coefficient between two traits. Usually this will be between 'his' and 'her' value of the same trait, i.e. an interspouse correlation r_{sp} but any pair of traits can be chosen.

Examples: (Everything typed by the user will be in bold type; anything else is therefore the computer replying or prompting for data or commands).

i) ORDER: **CORREL HIS HEIGHT V HER HEIGHT** (to calculate r_{sp} for height).

ii) ORDER: **CORREL HIS HEIGHT V HIS HEIGHT2** (to assess the repeatability of the measurement of height over the year. Any trait suffixed '2' means the same trait measured a year later).

iii) ORDER: **CORREL HIS HEIGHT V HIS WEIGHT** (to calculate the intercorrelation).

The program replies with the value of r, its standard deviation, the number of couples, and the significance of the result. Occasionally instead it prints an error message, that the value "is not computable". This is either because the number of couples is too low ($n < 11$) or the correlation is suspiciously high ($r > 0.99$), results of that strength being considered almost too good to be true. In fact they can occur occasionally. Thus the computer's reply to commands i, ii, and iii above might be:

HIS HEIGHT V HER HEIGHT N = 214 R = 0.293 SD = 0.066 ****

HIS HEIGHT V HIS HEIGHT2 IS NOT COMPUTABLE

HIS HEIGHT V HIS WEIGHT N = 214 R = 0.483 SD = 0.060 ****

The correlation coefficient is calculated by

$$r_{sp} = \frac{\sum y_1 y_2 - (\sum y_1 \sum y_2)/n}{\sqrt{(\sum y_1^2 - (\sum y_1)^2/n)(\sum y_2^2 - (\sum y_2)^2/n)}}$$

(FORMULA 2.1)

where y_1 and y_2 are the measurements of traits 1 and 2 (Snedecor & Cochran 1967).

The standard deviation is given by

$$SD = \sqrt{\frac{1 - r^2}{n - 2}}$$

(FORMULA 2.2)

The significance of the result is given by $z = r/SD$

where $z < 1.96$ implies $p > 0.05$, not significant
$1.96 \leqslant z < 2.58$ implies $0.05 \geqslant p > 0.01$, print '*'
$2.58 \leqslant z < 3.29$ implies $0.01 \geqslant p > 0.001$, print '**'
$3.29 \leqslant z < 3.89$ implies $0.001 \geqslant p > 0.0001$, print '***'
$3.89 \leqslant z$ implies $0.0001 \geqslant p$, print '****'

Thus in using the command CORREL we assume that both traits are Normally distributed.

2. SPLIT: This command treats the traits as all or none, according to whether the individual values lie above some arbitrary threshold set by the user. A concordance is printed, giving the number and percentage of couples where neither had the trait (0-0), only the woman had it (0-1), only the man had it (1-0), and both had it (1-1). Since the test is "above the threshold", values exactly equal to the threshold would score as zero. To avoid ambiguity, the threshold will always be set to a value that no measurement could exactly equal. The program goes on to print a correlation coefficient on the basis of this "all or nothing" sorting. A short cut formula is used. Let a be the number of marriages where both have the trait. If m men and f women are positive for the trait, then the number of marriages expected between them is mf/n.

Then Formula 2.1 above simplifies to:

$$r = \frac{a - mf/n}{\sqrt{(m - m^2/n)\ (f - f^2/n)}} \qquad \text{(FORMULA 2.3)}$$

Although high values of the trait have been considered positive and low values negative, it is easy to demonstrate that the result would have been the same had it been the other way round. When n-a, n-m and n-f are substituted for a, m and f, the resulting formula quickly reduces to that above. Note that we are using parametric methods on traits with only the values '0' and '1', which may cause loss of accuracy. Therefore the value of χ^2 is also given, calculated from the 2 x 2 contingency table which the concordance represents. Thus for religious practice among newly weds:

ORDER: **SPLIT HIS CHURCH V HER CHURCH** 0.5 0.5

```
HIS CHURCH V HER CHURCH SPLIT  0.5 0.5
CONCORDANCE:         0-0        0-1        1-0        1-1
NUMBER +%:          69(62%)    23(21%)     5(5%)     14(13%)
CHI-SQUARED:                               DF = 1    X2 = 14.677
CORRELATION:                               N = 111   R =  0.389 SD=0.088****
```

3. <u>BIAS</u> Detects whether a digit bias exists in a trait. The program selects an appropriate scale (eg whole units for pulse, weight, etc., tenth-units for height, skinfold etc.) and displays the distribution of readings, e.g.:

ORDER: **BIAS HER HEIGHT**

```
HER HEIGHT
SCALE OF READING: 0.00 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90
NO OF ENTRIES:     88   0    37   0    0    51   0    38   0    0
```

Readings of height are supposed to be to the nearest ¼ inch but some bias towards whole inches is revealed in this sample. Awareness of this serves as a reminder to the observer to avoid digit preference.

4. <u>GRAPH</u>: Draws a histogram of the trait's distribution; also gives the mean, standard deviation, and range. This command is a useful check on both the Normality of the data and its accuracy, as a typing error in one of the input files would generally lead to an extreme outlying value.

```
ORDER: GRAPH HIS SCHOOL
HIS SCHOOL
N = 216     RANGE IS 0.00 TO 5.00
MEAN = 2.04    SD = 1.23

0.0: ++
1.0: FIFTYPLUS:FIFTYPLUS:++++++++
2.0: +++++++++ !++++++++++!+++++++++!++++
3.0: +++++++++! +++++++++!+++++++++ !+++++
4.0: +++++++++!+++ ++++++!+++++++++!++
5.0: +++++
```

As a guide to numbers '!' marks each ten, and 'FIFTYPLUS:' replaces each whole fifty to prevent printing running off the end of the page. The command fails with the message that the trait 'is not graphable' if less than three values are known.

5. FIND: Locates all the couples having a particular value of a trait. For example, suppose a graph of resting pulse showed that one man had a pulse of 126. This is quite likely to be an error.

```
ORDER: FIND HIS RESTP 126
HIS RESTP    TARGET  126.0
52      126.0
```

The value belongs to the man in couple number 52. By checking the original sheet filled in at the visit, it might be discovered that, for instance, the resting and exercise pulse had been transposed when encoding the data.

6. CHANGE: Alters a particular value of a trait. The user replies to prompts with the reference number of the couple, and the new value to be inserted. Type reference number zero when no more changes are needed to that particular trait. For example, to correct the mistake discovered by FIND above:

```
ORDER: CHANGE HIS RESTP

HIS RESTP
PAIRNO = 52
VALUE  = 84
PAIRNO = 0

ORDER: CHANGE HIS TESTP

HIS TESTP
PAIRNO = 52
VALUE  = 126
PAIRNO = 0
```

This change only lasts until the end of the program run. If the computer's data file is to be permanently altered (and in this example it should be), it has to be separately edited.

7. CORRECT: Using the method of partial correlations, corrects the interspouse correlation for trait A for that existing for trait B. For example, there are correlations for smoking, and also for education. But smoking and education are linked, so perhaps the correlation for smoking is an artefact. The program calculates the matrix of all possible correlations between his smoking, her smoking, his education, her education, and tabulates these values. It then calculates the correlation for smoking which remains when education has been taken into account. See Chap. XIII of Snedecor & Cochran (1967), see also Chap. X. of Emery (1976) for a worked example. Of course the possibility might then arise that it is smoking that is genuinely correlated, and the similarity for education that is the artefact. To anticipate this question the computer goes on to calculate the correlation for education which remains when smoking has been taken into account:

ORDER: **CORRECT SMOKING V SCHOOL**

1) CORREL MATRIX SMOKING V SCHOOL

1.000	0.316	-0.376	-0.291
0.316	1.000	-0.271	-0.254
-0.376	-0.271	1.000	0.591
-0.291	-0.254	0.591	1.000

2) 1ST PARTCORRECTED FOR 2ND N = 111 R = 0.231 SD = 0.094*

3) 2ND PARTCORRECTED FOR 1ST N = 111 R = 0.528 SD = 0.082****

So in this example a correction for education has produced quite a marked reduction in the correlation for smoking; nevertheless it remains significant.

The matrix is displayed, but the partial correction 'is not computable' if either some matrix entry 'is not computable' as defined under 'CORREL', or the overall value of n (defined as the lowest n of any of the matrix entries filled) is less than fifty. The method of partial correlations is rather more sensitive to non-Normality than the method of calculating r so the conditions have to be more stringent. Also, the method is invalid if r has a very different value in different subgroups of the sample. One would need to check that, say, poorly educated people had the same correlation for smoking as well educated people; also that heavy smokers had the same correlation for education as light/non-smokers. This would be done using the command 'SELECT' which is described later.

8. CREATE: Generates a new trait which is a compound of two measured traits, and gives it a name for future reference (TRAIT1, TRAIT2, TRAIT3, or TRAIT4 - so that up to four 'created' traits may coexist). This would usually be the difference between two traits, e.g.:

ORDER: **CREATE HIS AGE - HER AGE 1**

HIS AGE - HER AGE IS THEIR TRAIT1

One could now see if a large difference in age was correlated with anything else, or graph the difference. (The GRAPH routine would ask for the scale to be given, as this can not be predicted for created traits.) Note that traits which can not be described as 'his' or 'her' are qualified as 'their' trait. Other examples are 'their pairno', the reference number, and 'their ascer', the method of ascertainment.

9. XTAB: Cross tabulates. In effect this is a two-dimensional version of GRAPH so that one can see who married whom with regard to a particular trait. For both dimensions, three numbers are required: the lowest value of interest, the scale size, and the highest value, in that order. This allows one to focus on any part of the distribution that is of particular interest. The example below cross tabulates social classes. GRAPH displays manual and non-manual workers together under '3' but for the present exercise

we wish to separate them. The scale size should therefore be 0.5. The program then goes on to give the nonparametric correlation coefficient rs, calculated by Formula 2.4:

ORDER: **XTAB HIS BESTJOB V HER BESTJOB 1 0.5 5 1 0.5 5**

```
HIS BESTJOB V HER BESTJOB 1 0.5 5  1  0.5 5

2ND VAR    ?? LO V4 ->             ->  V6    HJ
1ST VAR
V1                5   19
                  1   32   32   6   2
                       7   14   1   5
                      16   30  11  10   4
                       1    7   1   6   2
V3                     3    1
SPEARMANS RANK CORRELATION     N = 216     RS = 0.481  SD = 0.060 ***
```

Column headings (line beginning '2ND VAR') relate to the second variable: '??' means unknown values of 'her bestjob', 'V4' is the lowest value of interest (the fourth value typed in with the command), and 'V6' is the highest (the sixth value typed). 'LO' means values below V4 and 'HJ' values above V6; there aren't any here. Line headings (Column headed '1ST VAR') relate to the first variable. 'V1' is the lowest value, 'V3' the highest, and 'HI' are greater than V3. LO and ?? mean low and unknown as before. But to improve legibility all zero elements are suppressed, and several lines (??, LO, and HI for the first variable) are omitted completely if entirely zero. An error message appears ('CROSS TABLE MALFORMED') if in either dimension, i) the number of steps from low to high is less than 5, or greater than 13; ii) the low value is not actually lower than the high value.

10. SELECT: Chooses a particular subset of the sample. This may be either i) the entire group - in effect, cancelling a previous use of select since initially the entire group are being analysed, by default.

ORDER: **SELECT ALL**

or ii) only the group who are now to be defined, regardless of those who were analysed before, e.g. male nonsmokers.

ORDER: **SELECT ONLY HIS SMOKING = 0**

or iii) the group previously analysed, but plus or minus extra categories, e.g. having analysed male nonsmokers we now want to include those who had given up smoking by the second visit:

ORDER: **SELECT ABOVE OR HIS SMOKING2 = 0**

Compound statements may be used but it must be remembered that OR and AND are always used in their logical sense of set union and set intersection. Thus

ORDER: **SELECT ONLY HIS SMOKING2 = 0** AND HIS SMOKING = 0

would include only couples where the man was a non-smoker at <u>both</u> visits. The symbol '/' used before '<', '>' and '=' means 'NOT' and double negatives cancel each other. Thus, to analyse couples who got engaged:

ORDER: **SELECT ONLY THEIR ENG / = 0**

Occasionally it happens that 'SELECT' defines exactly the same group that was being analysed before, in which case the program prints 'no change made'. Otherwise it prints a line of asterisks across the page so that results relating to different subsamples are clearly separated, and then states how many couples are included. E.g. to study only first marriages:

ORDER: **SELECT ONLY HIS STATUS = 0 AND HER STATUS = 0**

```
SELECT ONLY    HIS STATUS = 0  AND  HER STATUS = 0
*********************************************************
156 COUPLES ANALYSED
```

Note that the command is echoed by the computer. This is unnecessary when the input and output devices are identical (teletype or VDU). But when output is going elsewhere (eg lineprinter) the results could be confusing unless this and other command routines made clear, as part of the output, precisely what they had been asked to do.

It often happens that SELECT results in a very small (or even empty) subgroup from which no useful information could be extracted. For example, selecting engaged couples who were heavy smokers would gather only four. The program would then waste time trying to perform a series of correlations only to find that in each case the result was 'not computable'. Therefore any 'selection' that leads to a subgroup of less than eleven is followed by the message 'SAMPLE INQUORATE'. The program ignores all further commands until it finds the next 'SELECT', with the exception that it will obey 'LIST' and 'FINISH' described below.

11. LIST: Gives the reference numbers of all the couples currently being analysed. Generally this is used after SELECT to see just who has been selected. E.g. after selecting engaged couples who are heavy smokers:

ORDER: **LIST**
LIST: 15 26 68 69

12. TTEST: Performs the t-test for matched pairs to investigate whether two traits have significantly different distributions.

ORDER: **TTEST HER BESTWT V HER BESTWT2**
HER BESTWT V HER BESTWT2 DF = 98 T = -1.411 SD = 10.040

The rise in weight in the women was not significant.

13. NONPARM: Calculates Spearmans rank correlation coefficient r_s. Some of the social traits are of dubious normality and it is useful to be able to check the results with a nonparametric method. The calculation is simple in principle but tedious in practice because one must rank and sort all the measurements of the two traits involved. The coefficient r is then derived from the sum of the squares of the differences in rank between each pair of measurements, with a correction for tied ranks. The following rapid method was therefore devised:

i) Use the routine XTAB to generate a table which is, in effect, a ranking of the measurements of both traits.

ii) For each row and column assign a mean ranking, MRANK and calculate T, the necessary correction for the tied ranks involved.

iii) For each table entry XTAB (I,J) summate $\sum d^2$ from (MRANK (Row I) - MRANK (Col J)) $*$ XTAB (I,J)

iv) Then $rs = \dfrac{\sum TI + \sum TJ - \sum d^2}{2\sqrt{\sum TI * \sum TJ}}$ (FORMULA 2.4)

where TI and TJ are corrections for tied ranks in rows I and columns J respectively. (Adapted from Siegal 1956)

The scale limits still have to be given as in XTAB, even though the table itself is not printed. Thus one would use the command XTAB when it was the display that was of main interest, with r_s an added bonus; and command NONPARM when there were a lot of correlations to be worked through since it is print-out, not arithmetic that slows down the program. Eg:

```
ORDER: NONPARM HIS BESTJOB V HER BESTJOB 1 0.5 5 1 0.5 5
SPEARMANS RANK CORRELATION      N = 216     RS = 0.481     SD = 0.060****
```

If $r_s > 0.99$ or $n < 11$ then r_s, like r, is marked 'not computable', although this method is probably accurate down to n = 4.

14. FINISH: Stops the program. If 15 or more correlations have been calculated during the session a warning is printed before signing off.

```
ORDER: FINISH
** WARNING: 63 CORRELATIONS COMPUTED
** SO 3.1 SIG RESULTS EXPECTED BY CHANCE
```

CHAPTER 3.

General Traits.

I. <u>Introduction</u>

II. <u>Methods</u>

a) Age
b) Fertility
c) Sibs
d) Status
e) Miscellaneous

III. <u>Results and Discussion</u>

a) Age
b) Fertility
c) Sibs
d) Status
e) Miscellaneous

CHAPTER THREE

I. INTRODUCTION

This chapter deals with the general biosocial traits of age, fertility, sibs, and marital status, as well as several miscellaneous traits not covered elsewhere. In the case of several of these traits it was not planned that correlations would be calculated, but they would simply be used as 'handles' by the computer program to select various subgroups.

II. METHODS

a) AGE

For each couple we know the dates of birth, of the two visits made (DATESEEN and DATESEEN2), the date married or moved in together (DATEWED and DATEMI) and of engagement (DATEENG) for ENG couples only. From these were calculated AGENOW, the age of each individual when first seen, and AGE, the age when married or moved in together. Thus for the engaged couples, AGE was greater than AGENOW. For ease of arithmetic dates were input as six-figure numbers and converted to decimal years. Thus 1st Oct 1978 is 011078 and is converted to 78.75. The precise formula was

$$\text{TIME} = \text{YEAR} + \frac{(\text{month of the year} - 1)}{12} + \frac{(\text{day of month} - 1)}{365}$$

but when the month is March through July then TIME := TIME $-\frac{1}{365}$(one day) and if the year exceeds 85 then TIME := TIME - 100 (one century). This formula is accurate to within one day throughout the year and no correction is needed for leap years. Note that 311285 represents the last day of 1985 but 010186 is the first day of 1886. This is to allow for the people who married in 1972 who were born in the last century.

b) FERTILITY.

Any children by previous marriages were coded as OLDKIDS, by the present marriage as KIDS, and future intended children as FUTKIDS. PREG equalled 1 when the woman was pregnant. Then TOTKIDS = OLDKIDS/X + KIDS + FUTKIDS but excluded PREG; likewise TOTKIDS2 = OLDKIDS/X + KIDS2 + FUTKIDS2 and ignored PREG2. The 'X' is because oldkids were only counted half when they resulted from a divorce.

c) SIBS

The traits SIBS, MATSIBS and PATSIBS give the number of full sibs of the person and his/her mother and father. TOTSIBS was the sum of all three traits.

d) STATUS

The STATUS was the person's marital status before the present marriage. (If the present marriage had been included the engaged couples would not have been comparable with the rest.) '0' represented bachelor or spinster, '1' widow or widower, '2' divorced once, '3' divorced at least once and widowed or divorced a further time.

Similarly PATSTAT and MATSTAT gave the parents' status. Note however that this could be ambiguous. Thus HIS PATSTAT=2 might mean that the man was born to his father's second marriage, the first having ended in divorce; or it could mean that the man's parents had split up. If also HIS MATSTAT=0 or 1 it must be the first possibility; but if HIS MATSTAT=2 it is probably (but not certainly) the second. The ambiguity was acceptable in the computer records since these were used only for preliminary rapid analysis before going back to the original handwritten records.

e) MISCELLANEOUS

This covers all the various other traits.

PAIRNO was the record number given to each couple; it was unique within the cohort but not within the sample (thus there was only one engaged couple number 1, but there was also a new '1' and an old '1').

ASCER was the method of ascertainment: 'E' for engagement notices, 'R' for registry office notices, 'W' for wedding notices, 'Y' for both 'R' and 'W'; 'O' for marriage lists, 'X' for mistakes from the 1972 group.

GROUP was the group into which the couple finally were placed: 'E' for engaged, 'N' for new, 'O' for old and 'V' for very old. Details of numbers, ascertainment and groups were given in Chapter 2.

ENG was (in new and old couples) whether the couple had been formally engaged. ENG=0 means not engaged. ENG=3 means engaged and announced the fact in a newspaper: therefore these couples were comparable with the engaged cohort. ENG=2 means engaged, but not announced in the papers. ENG=1 is a rather in-between category which had to be included because some people seemed to regard 'engagement' as a civil state in its own right intermediate between 'single' and 'married'. They definitely regarded themselves as having been engaged but this was not in the traditional sense of buying a ring or fixing a date to marry. This category also includes 'retrospective' engagements: although the couple had announced their 'engagement' they were already cohabiting before doing so.

In the engaged cohort where by definition they were in category '3', the term ENG was not applicable.

FATE was the fate of the partnership a year later. '5' meant still together and revisited. '0' meant not revisited but believed to be still together. '-3' meant one partner had died and '-5' meant the couple had split up.

III. RESULTS AND DISCUSSION

a) AGE

Table 3.1 shows a high correlation for age among married couples ($r \gtrsim .9$). It does not matter whether we take age at marriage, or when visited, as the criterion. However the engaged couples showed a much lower ($r = 454 \pm .110$) correlation. This is because the sample contained several couples with marked age differences, and one immediately wondered if these couples would fail to marry. In fact, most have already married and been revisited. As Table 2.1 shows, two split up, some were lost to follow up, and one engaged man died suddenly. Even if we only include couples who have married and been revisited, the correlation remains relatively low ($r = .512 \pm .121$).

Perhaps there is some hidden difference between the cohorts. Coleman (1977) found a very marked social class difference in assortative marriage for age. When the man was in classes I, IV or V the association was high (I r=.821, IV .761, V .752). But when the man was

TABLE 3.1

Correlations For Age

	ENG (n=68)		NEW (n=113)		OLD		
	r	SD	r	SD	n	r	SD
his age v her age	.454	.110	.903	.041	222	.888	.031
his agenow v her agenow	.501	.107	.906	.040	222	.910	.028
SUBJECTS OF COHORT							
his bestjob = I or II					104	.856	.052
his bestjob = IIIN or IIIM					98	.929	.038
his bestjob = IV or V					20	.856	.123
his education < O levels					109	.902	.042
his education = O/H/A levels					74	.888	.055
his education = university					39	.860	.084
his neuroticism 0 - 3					131	.904	.038
his neuroticism 3.5 - 6					90	.845	.058
his extroversion 0 - 3					97	.870	.051
his extroversion 3.5 - 6					114	.898	.042
male smokers					115	.834	.054
religious men					65	.930	.046
eng = 3					56	.837	.074
B x S marriages					173	.722	.053
him or her D or W					49	.875	.071
his age > 40					27	.740	.135

All correlations significant with $p < .0001$****

B = bachelor S = Spinster D = Divorced W = widowed 'bestjob' is social class as defined in Chapter 6. For details of neuroticism/extroversion scoring see Chapter 5.

middle class the association was lower (II r=.587, IIIN .350, IIIM .587). Futhermore, the association was much higher in remarriages (r=.901) than among first marriages (r=.648). Our engaged cohort was predominantly composed of people in classes II and IIIN and all but two were first marriages, the couples expected to have the lowest correlation.

Table 3.1 was therefore extended to show correlations for age in various subgroups. Many of these subgroups were almost unrepresented among the engaged (e.g. remarriages, low social classes) and therefore the old cohort had to be used. However this old cohort should be comparable to Coleman's sample, because his couples were also ascertained from marriages in 1972/73. The most striking feature of the tabulation was the constancy of the correlation shown, whether the subgroup was defined by class, education, personality, or whatever. In all cases r $\approx$ 0.8 even in those old couples where ENG = 3, ie those who would have formed the engaged cohort had this study been carried out eight years ago. Coleman's findings in Reading appear not to apply in Edinburgh, which leaves us without an explanation for the low correlation in the engaged. That correlation was only low compared to the 90% value in married couples, of course. Compared to the correlation for any other trait it was very high.

We may speculate that part of the discrepancy may lie in a secular change. People marry for the first time at all ages , and perhaps until a few years ago a high percentage of them would announce an engagement. It may be that nowadays only younger couples are formally engaged. There were no middle-aged or elderly couples among the present engaged cohort (unlike the 1972 'engaged' comparison group) and this would eliminate much of the variance of age. Intra-couple correlations can only be measured if there are inter-couple differences to set them against. And it is the young who are subjected to the full force of pro-engagement advertising, in cinema trailers etcetera. In fact it is tempting to conclude that it is not that gold rings are sold to seal an engagement, but that engagement is sold in order to provide a market for gold rings.

It is noteworthy that the lowest correlation among old couples in Table 3.1 was for first marriages: $r = 0.721 \pm .053$, rather higher than Coleman's 0.648 but clearly less than our overall 0.889. This merited further investigation. The correlation for age at marriage was therefore calculated for the entire marrying population of Scotland for the two relevant years of 1972 and 1977. The Registrar General for Scotland gives cross tabulations for age in Table Q2.7 (not reproduced here) in each annual report. Table Q2.7 was therefore examined for 1972 and 1977. There were separate tabulations for all marriages and for first marriages, so information about remarriages was quickly obtained by subtraction. The 1977 report also gave the same information for each local authority Region. The city of Edinburgh is in Lothian Region, and in 1977 accounted for 3591 of the 5441 marriages taking place in the Lothians. It was therefore quite straightforward in principle (though in practice a little tedious) to calculate the correlation for age from Table Q2.7.

The results have been listed in Table 3.2. We can confirm that remarriages were quite consistently (and significantly) better matched for age than first marriages, though not by the wide margin found by Coleman. Comparison of Tables 3.1 and 3.2 showed that the results from the present Edinburgh sample were in close agreement with those of all Scotland. It would be interesting to know if the results from Reading were typical of those for all England and Wales.

This closer matching for age among remarriages is a little surprising as one can immediately see two factors that militate against it. Firstly, people contemplating a remarriage are in a comparatively restricted market. There are few people of their own age available. One would predict, and the higher correlation confirmed, that they seldom find new partners among the bulk of unmarried youth. They must

TABLE 3.2

Correlation for age at marriage in the general population

B = bachelor S = Spinster D = divorced W = widowed
Thus couples were classified as remarriages (D or W) if <u>either</u> partner was previously divorced or widowed.

AREA	YEAR	TYPE	n	r	SD	
Scotland	1972	all	42139	.8804	.0023	****
"	"	B x S	35846	.7275	.0036	****
"	"	D or W	6293	.8473	.0068	****
Scotland	1977	all	37288	.8770	.0025	****
"	"	B x S	28528	.7093	.0042	****
"	"	D or W	8760	.8334	.0059	****
Lothian	1977	all	5441	.8793	.0065	****
"	"	B x S	3906	.703	.011	****
"	"	D or W	1535	.825	.014	****

look instead among the limited number of divorcees, widows, and ageing bachelors and spinsters, and it may well be that they have Hobson's choice. Yet we recall the high correlation ($r = 0.99$) among the Ramah Navaho Apaches (Spuhler 1968). Perhaps, given a scant choice, people match for age and all else goes by the board.

Secondly, it is likely that assortative marriage for age is lower in older people. That was the subjective impression formed during the study and later this was confirmed by figures to be given shortly. Of all the couples visited, there were more with a large age difference among the elderly (who were not necessarily members of the old cohort). Of course a difference of ten years or so is a much smaller proportion of a sixty year old's lifespan than of a thirty year old's.

The impression was confirmed at the foot of Table 3.1 : marriages where the man is over forty showed considerably lower correlations for age. One wonders whether this trend might be even more marked at greater ages, and whether it is true of both first and repeat marriages. The numbers in the present study did not suffice to analyse this, but again it was straightforward to extract this information from the RG's table Q2.7. Table 3.3 shows the correlations found for a range of husbands' ages. Values of n and SD are not listed but correlations have been given to four decimal places whenever SD < 0.01 and omitted completely when SD > 0.1 (implying $n < 100$). There was a striking decline in the correlation as the cut-off age was set successively higher. The trend was seen in first marriages and remarriages alike. The values for first marriages were always lower than for remarriages.

The lower correlation among the older is easier to explain than that of first marriages. People age at very different rates. It is not so noticeable before age 30, but thereafter divergence becomes evident. A crop of 50 year olds will vary widely in their physical fitness, attractiveness, libido, mental agility and general zest for life. One

TABLE 3.3

Correlation for age in older couples.

B x S : first marriages only   D/W : <u>either</u> partner previously divorced or widowed

	SCOTLAND 1972			SCOTLAND 1977			LOTHIAN 1977		
	ALL	B x S	D/W	ALL	B x S	D/W	ALL	B x S	D/W
all ages	.8804	.7275	.8473	.8770	.7092	.8334	.8793	.703	.825
man $\geqslant$ 30	.8165	.638	.8050	.8067	.634	.7980	.802	.627	.787
man $\geqslant$ 35	.7636	.571	.759	.7553	.599	.753	.755	.619	.748
man $\geqslant$ 40	.709	.522	.708	.695	.475	.698	.694		.692
man $\geqslant$ 45	.648	.455	.649	.638	.321	.644	.650		.648
man $\geqslant$ 50	.580	.371	.584	.572		.575	.597		.592
man $\geqslant$ 55	.507		.497	.508		.507	.499		.498

r given to four figures whenever SD $<$ 0.01, not printed if SD $>$ 0.1

would expect people to assort on the basis of this functional age, not by calendar years. Indeed the divergence may have precipitated the demise of the previous partnership. But proof is lacking, and it will be seen in the next chapter that there is no divergence in traits one might suppose to be markers of physical ageing (pulse, blood pressure, relative weight). Also, as Bowerman (1953) has shrewdly pointed out, we can only study the choices of those who liked what was available. Those who did not, or were not to the taste of others, remained unmarried and never appeared in the records. Presumably the more world-weary specimens are less likely to remarry, and even if they did would be reluctant to participate in the present survey.

The degree of homogamy for age found in Edinburgh was in line with results elsewhere (summarised by Roberts 1977). It is possible that it will rise in future, as remarriages form a greater proportion of the whole, and divorcees are freed for remarriage younger.

The unusual result of Pomerat (1936) was mentioned earlier. If lack of homogamy for age was generally associated with infertility that would be a very important finding, but is it true? The present sample contains several involuntarily infertile couples but not enough for investigation. Therefore, the correlation for age at marriage was calculated for one hundred consecutive infertile couples, from data kindly supplied by Dr. M. Lees of the RIE Infertility Clinic. The result is $r = .669 \pm .075$, a value typical of the general population. Infertility is not associated with low homogamy for age.

A number of traits were intercorrelated with age at marriage (AGE) or age when visited (AGE NOW). We will defer discussion to the sections on the relevant traits, but should at this stage consider the distinction between the two 'ages'. If a trait reflects either the ageing process (pulse etc.) or the length of marriage (number of children already born, etc.) then it should be related to age now. But if it reflects a secular trend (total fertility, etc.) it will be related instead to age at marriage. The discrepancy between age and age now has three components - the year of marriage, how long it took to arrange a visit, and whether the couple had been cohabiting before marriage. The first two are fairly constant for each cohort, in fact are part of the experimental design. But the third introduces a variation which is only partly corrected for by reassignment to different cohorts.

Figure B plots the number of couples who were cohabiting before marriage. The terms ENG, NEW and OLD here refer to the group in which they were originally ascertained. 12/81 ENG, 50/98 NEW and 30/224 OLD were cohabiting. As old couples were not routinely asked after the first hundred visits their rate is better expressed as 21/100. Cohabiting couples fell into three categories.

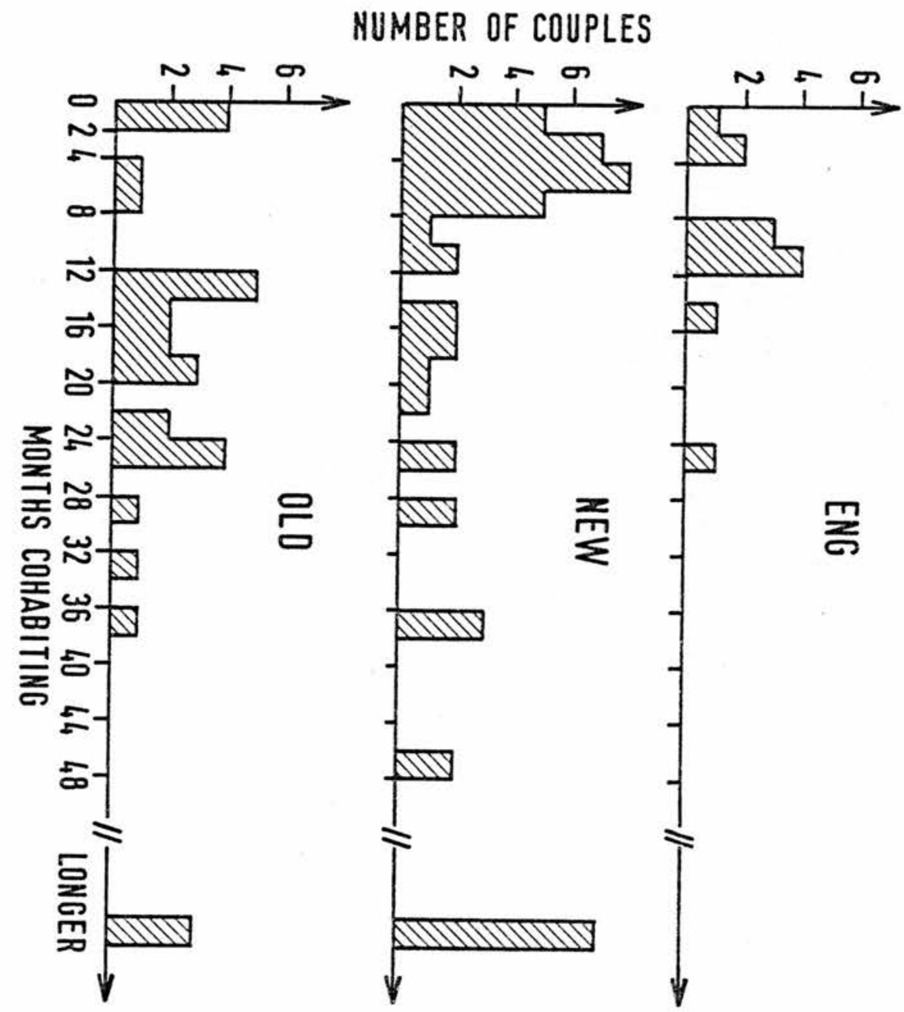
1) Young professional or semiprofessional couples, university educated, first marriage. It is quite common for university students to cohabit, and the student grant system strongly penalises married women. Many couples therefore postpone marriage until several months after graduation when their rising earnings mean there are tax advantages in marrying. Their careers may also draw them to opposite ends of the country, and if they move to a non-university and hence more conventional town they will find it more difficult to live together. Therefore as graduation approaches these couples must take stock and decide whether they really wish to stay together permanently. It would be interesting to know what proportion of these couples actually marry and how many split up.

2) Divorced people, older than average, all social classes. This group accounted for the very longest periods of cohabitation (up to 25 years). A major factor here was waiting for one partner's divorce to be finalised, so that they could marry. Recent divorce law reform (Divorce (Scotland) Act 1976) has facilitated this and probably in future such long waits will be rare.

FIGURE B

LENGTH OF TIME COHABITING BEFORE MARRIAGE

Cohorts 'Eng', 'New' and 'Old' here refer to the groups in which couples were originally ascertained, not those in which they were eventually analysed.



3) Very young, poorly educated, poor social class couples often without their own homes. Often the girl was pregnant before marriage and they were living with parents in cramped accommodation while trying to get their own Corporation house. It is already known that such marriages have a 50% failure rate (Dominian 1979c) and in fact 2 of the 7 such marriages had already broken up before the repeat visit a year later.

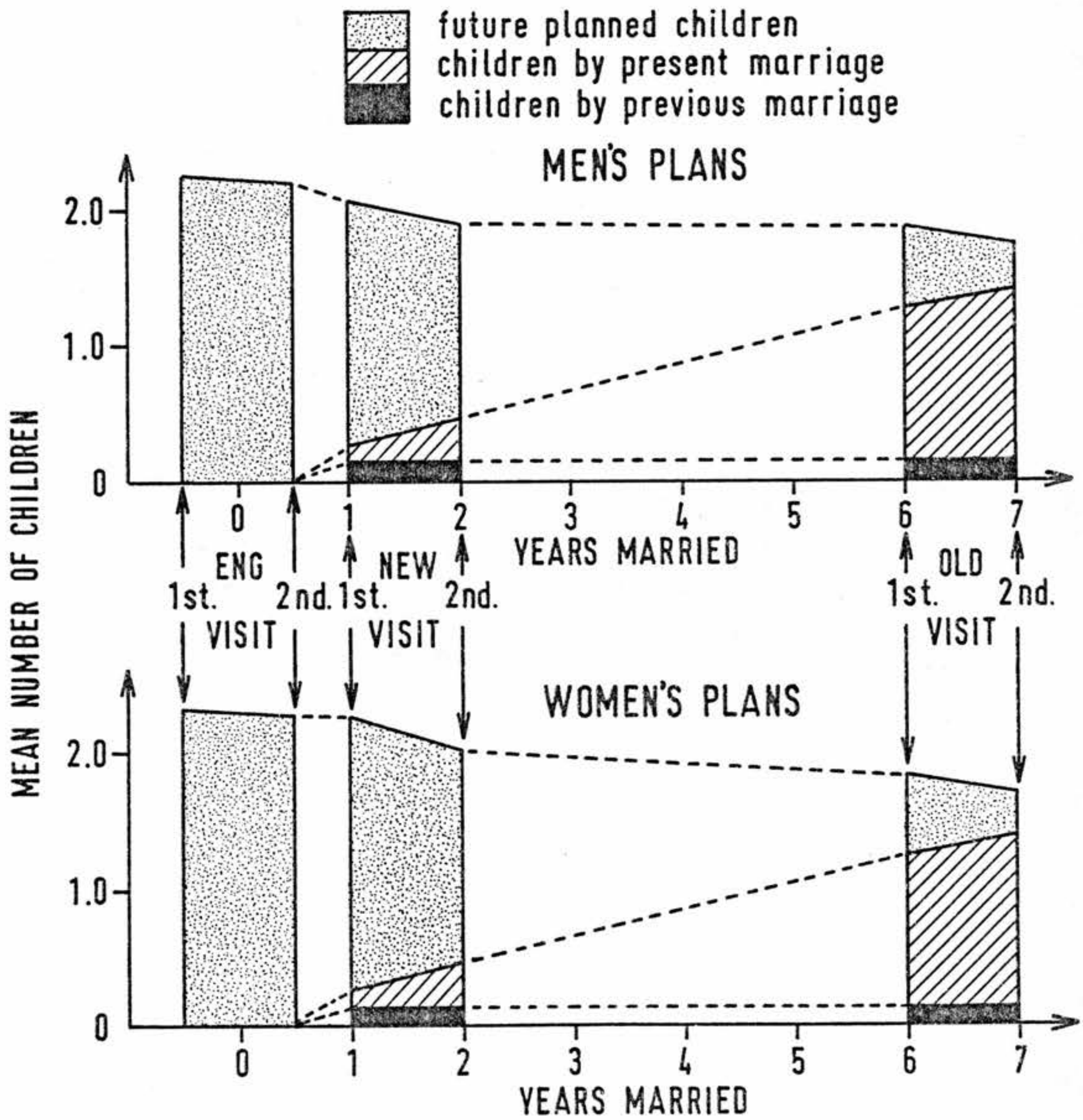
b) FERTILITY

Because most of the couples had only been married a few months or years their families were usually incomplete. Therefore it was only possible to ascertain their total eventual fertility from their planned family size. Since family planning advice and methods are so widely available these estimates should be reasonably accurate. Nevertheless they are only estimates and not the real thing. The total desired family size (TOTKIDS and TOTKIDS2) will underestimate fertility if there are a large number of 'accidents' i.e. unplanned pregnancies allowed to go to term. On the other hand, no correction was made for partnerships known to have split up.

Figure C illustrates how opinions on ideal family size changed during the study. The diagram shows total desired family (TOTKIDS and TOTKIDS2), and how the proportions shifted between children already born (KIDS, KIDS2), children planned for the future (FUTKIDS, FUTKIDS2) and children by previous marriages (OLDKIDS). When that previous marriage had ended in divorce the 'old' children only counted for a half-point each, on the Solomon principle. One could argue for a more refined assessment, e.g. taking account of who had custody, but it is unlikely that the extra complexity would have been justified.

The three cohorts have been plotted together in the same Figure in a way that implies that the newer will eventually evolve into the older. That is not strictly true because there are social and other differences in the composition of the cohorts. For instance the

FIGURE C: CHANGING FAMILY PLANS



engaged couples had no 'oldkids', and they did not acquire any when they married. Members of the engaged cohort who married are not the same as members of the newly-wed cohort.

The clearest trend (and one seen even within cohorts) was the general decline in fertility. As the cohorts aged and as the actual family size (old plus present children) rose, the total desired family size fell. The engaged couples were thinking in terms of two or three, which averaged out around 2.3. The newly weds wanted about two, though the women would have liked more. The older couples have settled for around 1.8 (i.e. the majority wanted two, but many also opted for nil or one), which is in line with other estimates of Scotland's population. Note that the estimates of total family size all had standard deviations of one, so it did not seem worthwhile to try subdividing the cohorts to discover whether changes in fertility were concentrated into any particular subgroup.

Table 3.4 gives correlations between desired family size (totkids and totkids2) and various social and general factors. The strongest correlation was with age and age now: the older the person, the less total family desired. Because this was seen in new and old cohorts it was unlikely to have been a secular trend or quirk of ascertainment. These correlations were not seen in the engaged, a group fairly homogeneous (but not necessarily homogamous) for age and marital status. Surprisingly there was no association with social class: there was no indication that working class people had larger families. But note that one source of such a correlation had been eliminated. If career-minded women were deferring having their family this would have reduced their fertility by lengthening the generation time, but the total eventual family size recorded as totkids would not have been altered. There was a suggestion that the better educated planned larger families. The correlation with education reached significance on three out of the twelve occasions tested.

TABLE 3.4

Correlation between total planned family size (totkids and totkids2) and other general and social traits.

	ENG		NEW		OLD	
HIS TOTKIDS V	(n=68)		(n=113)		(n=222)	
	r	SD	r	SD	r	SD
his age	-.018	.123	-.229	.092*	-.198	.066**
his agenow	.003	.123	-.240	.092**	-.177	.066**
his school	.218	.120	.204	.093*	.085	.067
their kids	NC		.025	.095	.532	.057****
his bestjob	.011	.123	-.089	.095	-.033	.067
his patjob	-.200	.122	-.126	.095	-.046	.068
HER TOTKIDS V						
her age	-.089	.123	-.204	.093*	-.416	.061****
her agenow	-.053	.123	-.212	.093*	-.381	.062****
her school	-.127	.122	.121	.094	.105	.067
their kids	NC		.032	.095	.569	.055****
her bestjob	.029	.123	-.107	.094	.005	.067
her patjob	.177	.121	.030	.095	-.008	.068
HIS TOTKIDS2 V	(n=52)		(n=97)		(n=198)	
his age	-.013	.141	-.177	.101	-.204	.070**
his agenow	-.010	.141	-.185	.101	-.171	.070*
his school	.308	.135*	.127	.102	.150	.070*
their kids	NC		.074	.102	.563	.059****
his bestjob	-.100	.141	.047	.102	-.086	.071
his patjob	-.194	.140	.006	.103	.025	.072
HER TOTKIDS2 V						
her age	-.027	.141	-.272	.098**	-.384	.066****
her agenow	-.007	.141	-.275	.098**	-.322	.068****
her school	.023	.141	.037	.101	.085	.071
their kids	NC		.020	.102	.646	.055****
her bestjob	.126	.140	-.035	.101	-.024	.071
her patjob	.258	.137	.065	.101	-.038	.072

'school' means education; 'bestjob' - social class; and 'patjob' - father's social class, as defined in Chapter 6.

'age' is age at marriage; 'agenow' age at first visit; 'kids' children already born to this partnership, as defined in the present chapter.

NC - not computable because no engaged couple had 'kids'

Couples' views on when they would wish to have their family were clearly moulded by the way the question was asked, and by the views of the partner - so although opinions were closely matched it would have been pointless to compute a correlation. Couples seldom pinned their decision to a single factor, but would identify one as weighing heavier than the others. Table 3.5 attempts to codify the answers to this open-ended question, although the bald percentages hardly do justice to the complex ifs and buts involved. These answers were based on the first visit, Those of the second visit have occasionally been used as a casting vote when someone was particularly unsure.

Money, house and job were all suggested in the question, yet compared to money, house and job appeared to be of minor importance. Obviously they are all linked: job implies money implies house, but few mentioned career prospects. The job was largely something to be endured for the sake of money. Money was seen to be less important in the older couples where it was more a case of "now or never". Those who had money either started their family or realised that they didn't want children and had just been using financial worries as an excuse. Those without money decided that, broke or not, they had better start if they wanted children.

Many also gave a time span, usually 1 to 5 years. These could be divided into couples without children (thus the time span related to the birth of the first child) and those considering addition to an existing family. In the former, the time span usually reflected confidence that everything else would be right. A person might say "In two years when" (not if) "Jim's salary's enough so I can stop working." Others mentioned having time alone together at the start of the marriage, and several were thinking of the woman's age. But those who already had children were more concerned about the last child's age: "Maybe when Fiona's around three, and then if it's another girl I might try once more for a boy".

TABLE 3.5

Family plans

% giving each reason as main deciding factor

	ENG		NEW		OLD	
	M	F	M	F	M	F
no more	7	6	17	17	64	65
1 - 5 years	53	54	25	30	11	9
money	23	28	20	12	6	4
now/en route	1	1	25	29	16	18
house/job	7	6	11	10	0	1
unsure	9	5	2	2	3	3

"Now/en route" includes pregnant women and those who had just delivered and would not have had time to reconsider their future. "No more" includes those who wanted never to have children, late marriages which would obviously be childless and those whose families were complete. Four were involuntarily infertile (including one man with cystic fibrosis) and four had curtailed their family because of genetic conditions (one malposition of great vessels, one Ehlers - Danlos syndrome, two neural tube defects). "Unsure" is a mixed category. It was possible to infer something about them from those who were unsure first time round but who later offered an opinion which allowed them to be categorised. Some may have been genuinely undecided, not only about timing but about whether to have any at all (and indeed, whether the marriage had a future). Others may have felt that family plans were something private that they did not wish to talk about. One or two were having difficulty, were beginning to suspect that they were infertile, and were rather defensive about the fact.

Table 3.6 shows a high correlation for total desired family size with $r > 0.5$ in all three cohorts. More interesting is the decline in planned family size seen in Figure C.

The impression was that vague optimistic plans for three or four had been whittled down to one or two by shortage of money, or by discovering that having children was not as idyllic as had been imagined. Although couples were asked about when and how many in realistic terms, not for what they might wish for in some ideal never-never land, a certain amount of wishful thinking was evident. For instance one man, asked when he would like a family, said "When we get a decent Tory Government to set the economy to rights". One year later a Tory government had been elected but he was having second thoughts. Several couples who at the first visit had hoped for a large family (the first being en route) next year declared emphatically "no more".

It is possible that family plans were being restricted by economic worries and fears about job security, and that family size might again expand given an improvement in the general economic climate.

c) SIBS

This was the only trait in the survey to be measured across three generations, in pat/matsibs, sibs and totkids. Table 3.6 showed little assortative marriage for sib number in the present generation (his sibs v her sibs r = 0.1 to 0.2). Likewise among the parents of old and new couples (matsibs v patsibs), r = 0.0 to 0.2. By contrast the parents of engaged couples show a very high correlation, 0.327 for men and 0.679 for women. Such correlations have previously been reported, but shown to be an artefact of secular trends in size of completed family (Beckmann & Elston 1962). The division of the sample by marriage cohort should have eliminated this artefact from the couples studied, but not necessarily from their parents. But the inconsistency between parents of married and engaged couples remained unaccounted for.

The correlation between pat/matsibs and sibs was seldom significant and there was no association between sibs and totkids (Table 3.7). Similarily there was a weak correlation between sibs and patjob but none between totkids and bestjob, nor between totkids and patjob (Table 3.4). Family size appeared not to be determined by social class, nor did those from the largest families want a large family for themselves. Perhaps we are witnessing the final cast-off of family plans from cultural ties, with attitudes to family size no longer being determined by family "imprinting". If this is true it may also be applicable to other traits such as voting habits and religious practice where traditionally one's family background has served as a model, but where other influences may have supplanted it.

TABLE 3.6

Sibs and fertility

	ENG			NEW			OLD		
	n	r	SD	n	r	SD	n	r	SD
his sibs v her sibs	68	.119	.122	113	.205	.093*	222	.084	.067
his patsibs v his matsibs	65	.327	.119**	109	.016	.097	215	.124	.068
her patsibs v her matsibs	68	.679	.090****	111	.184	.094	213	.083	.069
his totsibs v her totsibs	65	.484	.110****	107	.241	.095*	207	.086	.070
his totkids v her totkids	68	.505	.106****	113	.571	.078****	222	.612	.053****
his totkids2 v her totkids2	52	.750	.094****	97	.497	.089****	198	.747	.053****
his totkids v his totsibs	65	.320	.119**	109	.081	.096	215	-.063	.068
her totkids v her totsibs	68	.321	.117**	111	.030	.096	213	-.081	.069
his totkids v his totkids2	52	.754	.093****	97	.583	.083****	198	.838	.039****
her totkids v her totkids2	52	.690	.102****	99	.665	.076****	198	.794	.043****

TABLE 3.7

Number of sibs of each partner versus other social and general traits

	ENG (n=68)		NEW (n=113)		OLD (n=222)	
HIS SIBS V	r	SD	r	SD	r	SD
his age	-.241	.120*	.049	.095	.219	.066***
his agenow	-.240	.119*	.028	.095	.227	.066***
his matsibs	.086	.123	.042	.096	.030	.068
his patsibs	-.015	.126	.246	.093**	.168	.067*
his school	.125	.122	-.186	.093*	-.167	.066*
his bestjob	-.217	.120	.195	.093*	.174	.066**
his patjob	-.217	.121	.181	.094	.175	.067**
his totkids	.304	.117**	.028	.095	-.007	.067
HER SIBS V						
her age	-.236	.120*	-.162	.094	-.032	.067
her agenow	-.235	.120*	-.154	.094	.076	.067
her matsibs	.238	.120*	.193	.093*	.096	.067
her patsibs	.177	.121	.101	.095	.175	.068**
her school	-.133	.122	-.301	.091***	-.297	.064****
her bestjob	.149	.122	.271	.091**	.223	.066***
her patjob	-.077	.123	.204	.093*	.180	.067**
her totkids	.173	.121	.152	.093	.040	.067

d) MARITAL STATUS

The purpose of looking for correlations in status was to discover two things: whether divorce runs in families and whether divorcees etc. tend to pair off. But several factors combined to confuse the issue. Recent social changes made it impossible to compare former generations, so it was pointless to try to answer the first question. Secondly status was intercorrelated with age, which was highly homogamous. There was also the technical problem that parametric methods, such as partial correlations, were denied us in this case.

When marital status was simply cross tabulated (Table 3.10) the expected tendency to assortative marriage was seen. And we have already seen that correlation for age remained high even when the population was partitioned by status (Tables 3.1 and 3.2). The converse demonstration, that correlation for status remained after correction for age, was more involved. The analogous method was to partition the sample by age so finely that all correlation for age was purged, then see what assortment for status remained. But so strong was the association for age that the sample would have to be very thinly sliced indeed (five years at the maximum) and a very large sample would be needed. Bowerman (1953) found homogamy for status in a study of 37844 marriage records, Seattle 1939 - 1946. It was desirable to confirm this finding in a modern British population but the present 403 couples were insufficient for the purpose.

However the Registrar General's Tables Q2.5 and Q2.6 give information on status partitioned by age. Table 3.8 extracts some of the data, which applies to all Scottish marriages in 1977. Partitioning is by man's age at marriage, 30-65. Outside this range some combinations did not occur, causing a zero-element problem in calculation. Also, below age 30 a five-year slice would still contain assortment by age. Between ages 16 and 25 table Q2.7 gives ages by individual year, and the strong diagonal layout is visible even to the naked eye.

TABLE 3.8

Assortative marriage for status, partitioned by man's age

(All Scottish marriages 1977)

Agegroup of husband	Type of marriage BxS	BxW	BxD	WxS	WxW	WxD	DxS	DxW	DxD
All ages	28528	329	2242	330	597	399	2597	337	1929
30 - 34	1288	52	389	18	2	24	697	32	446
35 - 39	339	36	180	23	6	23	328	45	399
40 - 44	140	39	96	31	14	35	183	61	273
45 - 49	70	39	71	45	43	56	114	61	227
50 - 54	49	34	49	42	64	74	64	46	138
55 - 59	26	26	22	39	102	71	40	35	77
60 - 64	14	12	10	50	133	45	30	18	29
65 - 69	3	11	5	29	123	46	20	20	15

Bowerman had found that, for instance, the tendency of singletons (bachelors and spinsters) to prefer each other became most marked in the older couples as the field of eligibles shrank. His measure however was the "coefficient of contingency" which is bound to swell as the category dwindles. (This is discussed in more detail in Chap. 6, as Coleman has used the same measure in the analysis of social class.) A better measure is χ^2. Each age group gave a three by three table: B x S, B x W, etc. as in Table 3.8. For each of the three statuses, this was then reduced to a two by two table. For instance for divorced people the categories of interest would be D x D, D x other, other x D, and other x other. Then N = DD + DO + OD + OO

$$\text{and} \quad \frac{N.(\,|DD.OO - DO.OO| - N/2)^2}{(DD + DO).(DD + OD).(DO + OO).(OD + OO)}$$

is distributed as χ^2 with one degree of freedom (Chap. 8 of Snedecor & Cochran 1967). Results appear in Table 3.9.

B x S and D x D were strongly preferred among the younger people. This preference tailed off with age, and ceased to be significant around ages 60 - 65. Values of χ^2 for B x S paralleled those for D x D, though they were generally lower. Conversely, W x W only became a preferred combination at age 45, and thereafter tended to rise.

Why should there be assortative marriage for status? Other workers have spoken of "personality factors" but these seem rather nebulous. We are forced to speculate, but perhaps we can infer something from the very different behaviour of young and aged couples. In the young couples a major factor must be fertility. Bachelors and spinsters will prefer each other because they are unencumbered: by children, alimony or ex-spouses. Most singletons want a family, many remarrying people already have one and do not wish to add to it. A singleton who married a divorcee would then only be able to enjoy a family vicariously - at best as stepfather, at worst as honorary aunt. Do divorcees who have no children assort like singletons? This

TABLE 3.9

Assortative marriage for status : values of χ^2

	B x S	W x W	D x D
30 - 34	81.63	0.04	72.88
35 - 39	44.59	1.67	45.23
40 - 44	17.10	1.12	20.77
45 - 49	5.54	10.94	20.73
50 - 54	7.09	12.70	13.97
55 - 59	5.37	19.99	13.00
60 - 64	1.99	29.33	8.21
65 - 69	0.01	8.17	0.17

Values of χ^2 above 3.84 are significant with $p < 0.05$

Thus at age 30 - 34 both singletons and divorcees show significant assortative marriage but widows/widowers show no preference.

TABLE 3.10

Cross tabulation for status in present sample

B = bachelor S = Spinster D = Divorced W = Widowed

		ENG			NEW			OLD		
A. HUSBAND'S STATUS V WIFE'S STATUS										
HER STATUS:		S	W	D	S	W	D	S	W	D
HIS STATUS:	B	66			77	3	11	173	3	11
	W					2	1	4	7	2
	D	2			11		8	7	1	14
B. HUSBAND'S STATUS V HIS FATHER'S STATUS										
FATHER'S STATUS:		B	W	D	B	W	D	B	W	D
HIS STATUS:	B	58	1	7	79	4	8	165	6	16
	W				3			13		
	D	2			17	1	1	19	1	2
C. HUSBAND'S STATUS V HIS MOTHER'S STATUS										
MOTHER'S STATUS:		S	W	D	S	W	D	S	W	D
HIS STATUS:	B	62		4	82	2	7	166	5	16
	W				3			11	1	1
	D	1		1	17	1	1	19		3
D. WIFE'S STATUS V HER FATHER'S STATUS										
FATHER'S STATUS:		B	W	D	B	W	D	B	W	D
HER STATUS:	S	66		2	79		9	170	4	10
	W				5			9	1	1
	D				16		4	24		3
E. WIFE'S STATUS V HER MOTHER'S STATUS										
MOTHER'S STATUS:		S	W	D	S	W	D	S	W	D
HER STATUS:	S	65	1	2	78	2	8	168	8	8
	W				4	1		10		1
	D				16		4	23		4

Thus in Part A among engaged couples no woman had been married before; 66 of these spinsters married bachelors and 2 married divorced men.

information was not available. However, fertility can not have been the only factor. One would expect widowed people to have a fertility pattern similar to that of divorcees. Then W x D should also have been preferred. In effect, this would have inflated χ^2 for B x S and reduced it for D x D. D x D should have shown a pattern of values of χ^2 resembling that of W x W, not of B x S. This was not the case and therefore other factors are at work. One of these may be social class. The most vulnerable marriages are those of low social class adolescents. Their frequent early break-up will mean that there are many divorced people in the field of eligibles who have much in common besides their status.

In older couples fertility cannot be important. All are too old for child-bearing, and those who have had children have seen these grow up and leave home. (This is the third or "empty nest" stage defined by Dominian 1979g, but this staging is less helpful in the context of remarriage). Occupation is also less relevant because many will have retired already. The women's occupations will be especially misleading. Often they will have been housewives suddenly forced to take what work they can to earn a living. Neither sex is likely to have had the opportunity of higher education. This leaves personality factors, more or less by default. Other widowed people will form the bulk of the field of eligibles but a divorcee would probably be preferable to a singleton, as people who are still single at age 60+ have a high probability of being unsuitable in some way. These late remarriages, although they have zero fertility and therefore do not affect species fitness, have high survival value to the individuals concerned. Bereavement often precipitates illness and death of the remaining spouse (Young et al 1963) and the arrival of a new partner must have a profound influence on the will and ability to live.

All the divorced people in Tables 3.8 and 3.9 have been separated under the older legislation which required "grounds" for divorce such as desertion, adultery, and cruelty. (Divorces are still granted on these grounds where the newer conditions of separation for over two years do not apply). It will be some years before we can assess the effect of the new laws on remarriage patterns although one by-product - the ending of long periods of cohabitation - has already been mentioned.

e) MISCELLANEOUS TRAITS

At this stage we can briefly consider those traits which were included as padding, i.e. were never expected to be informative. These were name changes, consanguinity/isonymy, and handedness. Although there were a few name mutations induced by spelling mistakes on birth certificates, and changes associated with immigration and naturalisation, study of name changes was uninformative. Two couples found that they shared distant ancestors (beyond 5th degree relative) and the sample is therefore totally outbred. There were two isonymous marriages: Wilson x Wilson and Neave x Neave, both completely unrelated. There was no correlation for handedness (Overall value n = 338 r = .104 $\pm$ .055) which is fortunate as it would have been awkward to explain.

As expected, the sample was almost entirely ethnically homogenous, so no analysis of AM for race was possible. The only non-Europeans were two Iranian seamen (newly wed) and one Indian woman doctor (older group). They were simply pooled with the rest of the sample.

CHAPTER 4.

I. <u>Introduction</u>

II. <u>Methods</u>

a) Cardiac

b) Anthropometry

III. <u>Results</u>

a) Height

b) Weight & Skinfold Thickness

c) Pulse & Blood Pressure

IV. <u>Discussion</u>

CHAPTER FOUR

I. INTRODUCTION

It was shown earlier that anthropometric characteristics generally display weak assortative mating of the order $0.1 < r < 0.2$, but that possibly these traits were linked to height. Rather than pursue these often cryptic traits further, an attempt was made to understand AM for height itself in rather more detail. Also chosen for study were weight and skinfold thickness, pulse before and after exercise, and blood pressure.

II. METHODS

a) CARDIAC

Resting pulse was measured over 30 seconds, over 60 seconds if grossly irregular, with the subject sitting. Although by the time pulse was measured the couple had been sitting for several minutes, and were as near as possible at rest, this could not be entirely controlled. For instance, someone might have to go to answer the telephone. In these cases the resting pulse was repeated at the end of the interview (after the couples had answered the personality questionnaire) and the lower reading taken.

Readings were ignored if the person was on drugs such as digitalis, or had flu, or had been drinking heavily. However it would often be difficult to know whether a high pulse was due to drinking or smoking or whether it was the natural resting pulse. In any case it could be argued that if whenever a person was relaxing, he was smoking, then that was his 'natural' resting pulse. Therefore in all these doubtful cases the recorded pulse was included in the analysis, coded RESTP and RESTP2 from the first and second visits, but the compound trait BESTRESTP was taken as the lower of the two resting pulses. The rationale for this was that any artefact (smoking, flu, children starting to cry) would falsely elevate the pulse, but very few things would falsely lower it.

(Here and throughout the study the prefix BEST refers to a measurement which chooses the most accurate available data from several sources. It does not imply that it is a more desirable reading to have. As before, the suffix '2' means the same trait measured a year later.)

Blood pressure was measured by ausculation in the arm using an aneuroid sphygmomanometer. The fourth Korotkoff point was taken for the diastolic pressure. Blood pressure was seldom found to be elevated but in these few cases it was repeated later in the interview, and the lower value taken. The readings were ignored if the person was taking anti-hypertensive drugs, and omitted in several cases where it caused distress. The pressures were analysed under the headings SYSBP, DIABP, SYSBP2, and DIABP2 (i.e. Systolic and diastolic, first and second visits). As with resting pulse it seemed probable that any errors would be upwards (although drinking might lower the pressure slightly). Therefore BESTSYSBP was taken as the lower of the two systolics recorded, and BESTDIABP as the lower diastolic. Note that BESTSYSBP and BESTDIABP could refer to different years, e.g. if the first reading was 128/80 and the second 124/86.

Exercise pulse (TESTP and TESTP2) was measured after height and weight had been measured (see section below). It was clear that something like the McMaster step-test would be much too rigorous for many of the couples (indeed, might provoke a few coronaries) and that a less taxing version would have to be devised. The test instead was this: A standard, firm kitchen chair or dining room chair was used (i.e. the sort of chair that would be found in any home to be visited). The subject sat and held a weight in the outstretched arms - actually the bathroom scales, which weighed six pounds. He or she then stood up and sat down ten times, keeping the weight outstretched and level with the shoulders at all times. After sitting down for the tenth time the subject rested. After ten seconds recovery time (allowing the pulse to be found, and irregularities due to heavy breathing to be smoothed out), the pulse was recorded for fifteen seconds. The actions required were demonstrated beforehand, and the subjects prompted "faster!" or "slower!" so that exercise was completed in 20 seconds, plus or minus 5 seconds. If exercise could not be completed in 25 seconds the reading was disregarded. The test was not carried out on people who were frail, arthritic, or heavily pregnant.

BESTTESTP was taken as the lower of the two years' readings.

All pulses were encoded as a rate per minute, for ease of comparison. For each year, the ratio between the resting pulse and exercise pulse was calculated:

PRATIO = (TESTP/RESTP) * 100
PRATIO2 = (TESTP2/RESTP2) * 100

Likewise, the best available estimate of pulse ratio was calculated as:

BESTPRATIO = (BESTTESTP/BESTRESTP) * 100

b) ANTHROPOMETRY

Couples were first asked how tall they thought they were and how much they thought they weighed. Pregnant women were also asked how much they weighed before pregnancy. Only when this information had been recorded were the bathroom scales and tape measure revealed.

All recordings of height were in feet and inches, and all recordings of weight were in stones and pounds. At the beginning of the study it had been planned to make all measurements metric. This was quickly abandoned because people's estimates of their own height and weight were given in imperial units: this would have led to working in two sets of units and increased the chance of gross errors. Very occasionally somebody could only estimate his/her measurement in metric units. In this case the metric value was written down then the imperial equivalent worked out on the spot (both tape and scales had both sets of units). Of course the units do not affect the strength of the correlation, which is a dimensionless number.

At the second visit the couple would have known in advance that there were scales and tape. On this occasion only weight was estimated. The traits estimated were encoded as ESTHT, ESTWT, and ESTWT2.

Initially height was measured in two ways: by tape measure and by wooden measuring frame. It was believed that the frame would be an accurate standard against which the tape could be compared. In fact, the frame was found to be completely impracticable. There was difficulty in zeroing it, and considerable play in the horizontal bar, so that readings could be two inches out. It was therefore abandoned, with some relief as it was clumsy to transport. All readings of measured height (HEIGHT, HEIGHT2) were then carried out as follows.

A suitable mark on the wall (usually a lightswitch) was chosen, and its height measured. (It was not difficult to find a mark which would be an exact inch or half-inch high.) The subject then stood in front of the mark, in socks but without shoes, with heels together and against the wall, shoulders back against the wall but not leaning on it, head held horizontally. A straight edge (actually the box containing the skinfold calipers) was lowered onto the person's head and pressed against the wall. The distance between the bottom of the box and the original mark was then measured, and the height found to the nearest quarter-inch. From time to time the tape measure was compared against some other measure to confirm that it had not stretched with use.

The compound measurement BESTHT was worked out as follows: If the person's height had actually been measured on both visits (as in the vast majority of couples) then BESTHT was the mean of the two readings. If only one measurement was available, BESTHT was that value. Only if height had been measured on neither occasion (e.g. if a very arthritic person could not stand straight up) was BESTHT taken as the estimated height, ESTHT.

Initially weight was measured on two sets of bathroom scales, one set being the older but also rather lighter and smaller. It was quickly apparent that the older set were not acceptable because they were hard to zero, slow to settle down on the reading, and probably not robust enough to last out the study. The newer set, which were Salter scales, were therefore used for all couples.

With the scales on a firm surface, subjects were weighed wearing their ordinary indoor clothing but without shoes or any particularly heavy items such as jackets. It was pointed out to the couples that the scales could read over a stone heavier on a soft surface and this guaranteed that they found the firmest surface available. Readings were taken to the nearest pound. Although pregnant women were weighed these results were ignored, as were those of recently delivered women whose weight had not yet restabilised.

For each year, the best estimate of weight (BESTWT, BESTWT2) was taken as the actual weight measured if available, otherwise the estimated weight for that year. Changes in weight were less likely to be measurement errors (as with height), or ephemeral artefacts (as with pulse and blood pressure), so there seemed no reason to derive a 'best' measurement from two years' combined data.

Skinfold thickness was measured with Harpenden skinfold calipers. The sites measured were the left triceps, left biceps, chin and left knee. These sites were chosen because they were easily accessible, and often visible, in a clothed subject. Different sites have been used in most nutritional studies (subscapular, mid-axillary) but they are not readily accessible and here it was not intended to calculate total body fat. To measure the triceps skinfold, the left elbow was flexed, and the midpoint between olecranon and acromion noted. The skin was picked up in a fold an inch above this site, then the calipers applied to the spot. Two seconds were allowed for the needle to settle then the reading taken to the nearest 0.2mm (but readings over 20mm were only recorded to the nearest whole mm). The biceps reading was taken from the point on the arm diametrically opposite to the point of the triceps reading.

At the chin, the fold was not picked up with the fingers but the open jaw of the calipers rested on the skin $1\frac{1}{2}$" behind the tip of the chin. Then with the subject's head level the calipers were allowed to close and take in a fold of skin as they did so.

Behind the left knee, the fold was raised in the midline of the popliteal fossa, just below the flexion skin crease, with the subject standing and slightly flexing the joint. Occasionally because of skin-tethering the reading had to be taken up to an inch above or below this point.

Very occasionally a site was inaccessible behind clothing that could not conveniently be removed, and therefore the following compromise had to be accepted. The layers of clothing, but not skin, were picked up and measured, and the calipers rezeroed at this thickness. The clothing and skinfold were then picked up together and measured. (The thickness of a pair of tights was negligible and no correction was necessary for these. Nor was any correction made for the thickness of superficial skin overlaying the fat.)

The skinfolds were encoded as the measured values but the computer converted them all into logarithms to the base e, as it is known that skinfolds are log Normal (Durnin and Womersly 1974). Reference to these traits (TRISKF, TRISKF2, BISKF, BISKF2, CHINSKF, CHINSKF2 and LEGSKF, LEGSKF2) therefore concern the $\log_e$ values. In addition, total skinfold thickness (TOTSKF, TOTSKF2) was calculated as the sum of the year's four $\log_e$ values. BESTSKF was the average of these two totals.

III. RESULTS

a) HEIGHT

Table 4.1 allows us to conclude, firstly that there is quite clearly assortative marriage for height. The value of r is .301 for engaged, .170 for newly weds and .304 for old couples. The value for newly weds was not quite significant. Secondly, the method of measuring height was accurate. The repeatability over the year was at least 99%. Simply asking each person his or her height gave 95% agreement with the measured value, although four people were out by as much as four inches.

Table 4.2 shows the range, mean and standard deviation of the 'best' heights calculated. The men were some 5 to 5½ inches taller than the women and the engaged couples tended to be taller than the other groups. However there was a fair amount of overlap in the distributions.

A number of questions are raised by these results. Firstly, might the correlation for height have been inflated by some other variable, such as age? Although these couples were all adult, and were not going to grow any more, some were elderly and were quite visibly shrinking,

TABLE 4. 1

Correlations between various measures of height

	ENG			NEW			OLD		
	n	r	SD	n	r	SD	n	r	SD
his height v her height	68	.312	.117**	112	.167	.094	220	.306	.064 ****
his estht v his height	67	.963	.033****	113	.949	.030****	222	.957	.020 ****
her estht v her height	68	.968	.031****	111	.941	.032****	219	.935	.024 ****
his height v his height2	52	1	****	99	1	****	198	1	****
her height v her height2	52	1	****	99	.990	.015****	196	1	****
his bestht v her bestht	68	.301	.117*	113	.170	.094	222	.304	.064 ****
his estht v her estht	68	.292	.118*	112	.171	.094	221	.254	.065 ****

where 'estht' is estimated height and 'bestht' is best estimate from measurements (height, height2) or estht,
Correlations of 1 are those marked 'not computable' because r > 0.99

TABLE 4.2

'best' heights of couples (in inches)

	ENG	NEW	OLD
his bestht : range	65 - 77	$63\frac{1}{4}$ - 76	$60\frac{1}{4}$ - 76
mean & SD	69.79 ± 2.38	69.27 ± 2.47	68.82 ± 2.70
her bestht : range	$59\frac{5}{8}$ - $70\frac{1}{8}$	$57\frac{3}{4}$ - $68\frac{5}{8}$	57 - $70\frac{3}{8}$
mean & SD	64.28 ± 2.43	63.84 ± 2.27	63.67 ± 2.39

It was noticeable during the visits that the largest discrepancies between estimated and measured height occurred in elderly people who could no longer straighten their spines. The cross-correlations were therefore calculated between height and a selection of other traits: AGENOW (i.e. age when visited, not age at marriage), BESTJOB, education, neuroticism, extroversion, and BESTWT. These are presented in Table 4.3.

A strong intercorrelation with weight was seen and this was to be expected. As we know that height influences weight, but weight does not influence height, it seems more logical to discuss this connection in the section on weight.

The association between height and social class and education suggests a nutritional effect on stature. This was a commonplace finding in bygone years, before welfare state milk and orange juice; perhaps the effect is lingering on in the older couples. But as Wilson and Nias (1976) point out, tall people may be likelier to be promoted, and higher class people are perceived as taller; these two factors could become self-perpetuating. Table 4.3 also reveals the tendency to shrink with age: 'agenow' and height were negatively correlated especially in old couples. But no association was seen between stature and personality traits.

Table 4.10 shows that there was relatively little change in the correlation for height after correction for age, social class, and education. Nor was the correlation for these traits greatly influenced by correction for height. Assortative marriage for height therefore appeared to be independent of these other traits.

A second question is how people's estimates of their own heights compared with reality - did relatively short men exaggerate their height or did tall women underestimate it? Or did couples tend to estimate themselves as more similar than they really were? The last suggestion is easily refuted by Table 4.1 since the correlation between his and her 'estht' is close to that between his and her height.

TABLE 4.3

Intercorrelation of best heights with other traits

	ENG (n = 68)				NEW (n = 113)				OLD (n = 222)			
	M		F		M		F		M		F	
	r	SD	r	SD	r	SD	r	SD	r	SD	r	SD
age now	.022	.123	-.136	.122	-.202	.093*	-.120	.094	-.231	.066***	-.202	.066**
best job	-.196	.121	-.158	.122	-.129	.094	-.100	.094	-.200	.066**	-.142	.067*
education	-.056	.123	.176	.121	.179	.093	.233	.092*	.198	.066**	.153	.067*
neuro	-.046	.123	.122	.122	-.066	.095	-.076	.095	-.025	.068	-.083	.067
extro	-.080	.123	.071	.123	-.068	.095	-.184	.093	.025	.068	-.085	.067
bestwt	.480	.108****	.486	.108****	.374	.088****	.429	.086****	.515	.058****	.332	.064****
totkids2	-.124	.140	-.028	.141	-.007	.103	-.039	.101	.019	.071	-.005	.071

Thus column headed 'M' is 'his bestht v his age now', etc.; that headed 'F' is 'her bestht v her age now', etc.

TABLE 4.4

People's perceptions of their own heights

(excluding couples where his agenow > 45)

Definitions: trait1 = her height - his height
trait2 = his estht - his height
trait3 = her estht - her height

	ENG (n=68)		NEW (n=108)		OLD (n=182)	
trait1 v trait2	-.051	.123	.135	.096	.112	.074
trait1 v trait3	-.049	.123	-.008	.098	.117	.076
t2 v his height	.013	.123	-.046	.097	-.258	.072***
t3 v her height	-.077	.123	.009	.098	.092	.075

TABLE 4.5

Interspouse Correlation for height in different social strata

	ENG			NEW			OLD		
SUBSET	n	r	SD	n	r	SD	n	r	SD
his bestjob = I, II, or IIIN	58	.327	.126*	103	.200	.097*	132	.241	.085**
his bestjob = IIIM, IV or V		NC			NC		90	.331	.101***
his educ = Highers +	54	.229	.135	53	.058	.140	78	.295	.110**
his educ = 'O's or less	14	.438	.260	60	.239	.128	144	.266	.081***

NC: marked 'not computable', not enough couples in this subset.

TABLE 4.6

Interspouse correlation for various measures of weight

	ENG			NEW			OLD		
	n	r	SD	n	r	SD	n	r	SD
his weight v her weight	68	-.167	.121	98	.156	.101	193	.098	.072
his estwt v her estwt	68	-.164	.121	113	.184	.093*	221	.111	.067
his bestwt v her bestwt	68	-.167	.121	113	.154	.094	222	.120	.067
his weight v his estwt	68	.965	.032****	113	.969	.023****	221	.971	.016****
her weight v her estwt	68	.959	.035****	98	.978	.021****	193	.978	.015****
his weight2 v her weight	52	.035	.141	90	.281	.102**	179	.162	.074
his estwt2 v her estwt2	52	.014	.141	98	.242	.099*	197	.156	.071*
his bestwt2 v her bestwt2	52	.035	.141	94	.243	.098*	197	.172	.071*
his bestwt v his bestwt2	52	.935	.050****	99	.951	.031****	197	.962	.020****
her bestwt v her bestwt2	52	.852	.074****	99	.878	.049****	198	.942	.024****
his relwt v her relwt	68	-.084	.123	113	.178	.093	222	.092	.067
his relwt2 v her relwt2	52	.163	.140	99	.225	.099*	198	.099	.071

where 'weight' = measured weight, 'estwt' estimated weight; 'bestwt' is measured weight if available otherwise estwt
'relwt' is relative weight by Garrows formula.

To determine whether short men exaggerated or tall women underestimated it was first necessary to eliminate people over 45, as some of them had dwindled markedly from their original height. Then the difference between his and her height was defined as TRAIT1, the difference between his real and estimated height as TRAIT2, and the difference between her real and estimated height as TRAIT3. Then it is possible to see if the tendency to misestimate one's own height was correlated with either relative (their trait1) or absolute (his/her height) measures of stature. In fact, all these correlations were non-significant (Table 4.4). Therefore no consistent mis-estimation of height has been found.

Thirdly, was assortative marriage for height confined to one particular social group, as it was in Oxfordshire (Harrison et al 1976)? Table 4.5 shows the correlation for height in different social and educational strata. In fact there were insufficient class IV and V couples for calculation, but classes I, II and III, which in Harrison's study did not assort for height, here show a significant correlation. Moreover, the correlation was not greatly altered by subdivision by education. The difference may be that the Oxfordshire study was rural, and the present study is urban. In a rural population Classes IV and V are composed mainly of farm labourers. Their families have probably lived locally for many generations and they themselves are probably marrying the girl next door. Classes IV and V in cities are a more rootless, 'anomic' group.

Fourthly, why should there be assortative marriage for height? Is it the man or the woman who is exercising the choice or both? Here we are on less certain ground as the three possibilities are mathematically indistinguishable in the recorded data. However the impression was that it was the women who were actively avoiding shorter men. When, during the visits, it was mentioned that height was turning out to be an important factor the husband would appear noncommital but the wife would declare that she would have felt ill at ease with a shorter partner. Wilson and Nias (1976) say that "It is a cardinal rule of dating and mating that the man be taller than the woman, so relative height may be just as important as absolute height". This is certainly true of the present sample, even though we live in an age when many other "cardinal rules of dating and mating" may be ignored as passé. Although Table 4.2 shows a considerable overlap between male and female distribution of heights, in only four of the 403 couples was the woman as tall as, or taller than, the man.

A separate analysis was made of couples where the man was relatively short. To get a worthwhile number the criterion was relaxed to HIS HEIGHT - HER HEIGHT $< 2.2''$ by which 58 couples were included. A search was made for ways in which these couples might differ from the rest. However, correlations for age, social class, education and personality factors (not tabulated) were similar to those for the whole group. The correlation for height itself rose to $r = 0.8$ but that was an artefact. Since men are taller, selecting relatively shorter men was bound to result in a very closely matched subgroup.

b) WEIGHT AND SKINFOLD THICKNESS

Table 4.6 shows that people's estimates of their own weight were highly accurate in both sexes, i.e. the correlation 'weight v estwt' had values of r of .96 - .98. However, unlike most European populations previously studied, the present sample did not appear to be practising assortative marriage for weight. There is a suggestion in the table that there was a transient correlation shortly after marriage. The engaged couples had a negative but insignificant value; the newly-weds had a positive value which reached significance one year later; the older couples again had a nonsignificant value. This is similar to the finding of Spuhler (1967), that weight at marriage was correlated but that the correlation subsequently disappeared. However it seemed equally probable that this ephemeral similarity was a statistical artefact.

TABLE 4.7

Weight in pounds : mean and SD

Relwts as % of optimum

	ENG		NEW		OLD	
	mean	SD	mean	SD	mean	SD
his bestwt	160.59	18.73	160.65	21.31	161.44	22.47
her bestwt	122.97	14.36	126.36	17.39	134.43	23.37
his bestwt2	163.48	20.04	159.79	22.92	161.75	21.65
her bestwt2	124.04	12.36	129.46	20.82	134.68	22.03
his relwt	102.99	10.58	104.25	13.11	106.39	12.68
her relwt	97.39	9.69	101.22	12.26	107.99	17.24
his relwt2	104.31	11.69	103.96	13.93	106.50	12.61
her relwt2	97.91	7.64	103.36	15.13	107.97	16.98

Table 4.7 shows that the men's weight was remarkably constant, but the women's was steadily rising. This could be related to increasing length of marriage, or increasing age, or birth of children. As the rise was not actually a significant one, no calculations were done to decide this. The strong correlation between height and weight has already been noted and it therefore seemed worthwhile to study relative weight - i.e. weight taking account of height. Garrow (1979) found that the best measure is W/H^2, where W is weight in kilograms and H is height in metres. The ideal range (ie that associated with the lowest mortality) is 20 - 25 for men and 19 - 24 for women. For convenience, relative weight was here defined as

$$\text{RELWT} = (3124.4 \times \frac{\text{BESTWT}}{(\text{BESTHT})^2}) + 4.4 \text{ if female}$$

and likewise $$\text{RELWT2} = (3124.4 \times \frac{\text{BESTWT2}}{(\text{BESTHT})^2}) + 4.4 \text{ if female}$$

(note the difference between $(\text{bestht})^2$ i.e. best height squared and bestwt2 i.e. best weight at the second visit).

The above definition was chosen so that with weight in pounds and height in inches the optimum RELWT would be 100 in both males and females. The desirable range was then 90 to 110% of the optimum. For illustration Table 4.8 gives ideal weights for men and women of various heights, using Garrow's formula.

Table 4.6 shows that, as with weight, relative weight showed a transient similarity shortly after marriage. A tendency to increase in relative weight is also seen in Table 4.7. It is clearer than that for actual weight because the older couples were shorter, which diminished their actual weight but increased their relative weight. However, again it was nonsignificant and was not examined further.

Table 4.9 shows the correlations for skinfold thickness over triceps, biceps, chin and leg, and for the total of the four readings. Additionally BESTSKF has been calculated as the average of the two totals, totskf and totskf2. Little or no correlation was seen in engaged and newly-wed couples. But in older couples there was a strong correlation at the chin.

TABLE 4.8

Ideal weights by Garrow's formula

height (in)	man's wt (lb)	woman's wt (lb)
58	108	103
59	111	107
60	115	110
61	119	114
62	123	118
63	127	121
64	131	125
65	135	129
66	139	133
67	144	137
68	148	141
69	152	146
70	157	150
71	161	154
72	166	159
73	171	163
74	175	168
75	180	172

TABLE 4.9

Interspouse correlations for skinfold thickness

	ENG			NEW			OLD		
	n	r	SD	n	r	SD	n	r	SD
triskf	67	.000	.124	104	.087	.099	222	.002	.067
biskf	67	.296	.118*	104	.053	.099	222	.124	.067
chinskf	67	.061	.124	104	.075	.099	222	.238	.065***
legskf	67	.198	.122	104	.147	.098	222	.107	.067
totskf	67	.137	.123	104	.093	.099	222	.139	.067*
triskf2	52	.127	.140	98	.135	.101	196	.067	.072
biskf2	52	.142	.140	98	.105	.101	196	.129	.071
chinskf2	52	.071	.141	98	.200	.100*	196	.279	.069****
legskf2	52	.088	.141	98	.092	.102	196	.174	.071*
totskf2	52	.092	.141	98	.146	.101	196	.203	.070****
bestskf	67	.145	.123	110	.135	.095	222	.174	.066**

bestskf weighted mean r = $.157 \pm .051$**

TABLE 4.10

Correlation for trait A which remains when trait B is held constant

		ENG (n=68)		NEW (n=113)		OLD (n=222)	
trait A	trait B	r	SD	r	SD	r	SD
bestht	age now	.314	.119**	.165	.094	.285	.065****
age now	bestht	.487	.109****	.904	.041****	.906	.029****
bestht	bestjob	.265	.121*	.151	.095	.275	.065****
bestjob	bestht	.256	.121*	.541	.081****	.460	.060****
bestht	education	.306	.119*	.155	.095	.274	.065****
education	bestht	.475	.110****	.585	.078****	.623	.053****
sysbp	age now	.095	.125	.210	.095*	.100	.068
age now	sysbp	.505	.109****	.906	.041****	.872	.034****
chinskf	age now	.054	.126	.053	.100	.181	.067**
age now	chinskf	.496	.109****	.905	.042****	.901	.029****
besttestp	smoking	.323	.119**	.198	.095*	.008	.070
smoking	besttestp	.400	.115**	.264	.093**	.311	.066****

where : 'age now' is age when visited.
'bestht' is best estimate of height.
'bestjob' is best estimate of social class
'sysbp' is systolic blood pressure.
'chinskf' is skinfold thickness at chin.
'besttestp' is lower of the two years' exercise pulses testp and testp2

These couples also had a weak correlation for weight. Two explanations are possible. It may be that many older people acquire double chins and that the association was an artefact of matching for age, or it may be that certain couples both tend to grow double chins and this therefore represented convergence. One may therefore distinguish the two possibilities by performing partial correlations of chin skinfold with age, and this has been done in Table 4.10. The correlation for skinfold has been greatly reduced, from .24 to .18, but remains highly significant. This implies that much of the similarity may be an artefact of age but there is also evidence of genuine convergence.

The repeatability of the readings (tested by correlations of the form CORREL HIS LEGSKF V HIS LEGSKF2, etc. - not tabulated) was around 80%. Since the repeatability of weight was 90% or more there may have been a small unrepeatable factor in the readings over and above that due to change of weight. This was presumably an instrument/observer error, but the accuracy seemed acceptable.

c) PULSE AND BLOOD PRESSURE

Table 4.11 shows that any similarity for pulse (at rest, after exercise, and the ratio between the two) was slight. As with weight, there may have been a transient correlation about the time of marriage. The correlations for systolic pressure were nonsignificant in engaged couples (r= .099) but high in new (.218) and old (.258) couples, whilst diastolic pressure was never correlated (Table 4.12). Whilst it is conceivable that shared domestic environment could have produced such a convergence of systolic pressure (e.g. through the amount of salt consumed), it was likelier to be an artefact, because it had almost disappeared a year later. When correction was made for age now (Table 4.10) the association vanished. In the case of pulse, correction for age now had no effect but correction for smoking habits greatly reduced the correlation. Smith (1946) had found pulse to be positively correlated before exercise but negatively afterwards, while blood pressure was not correlated.

TABLE 4.11

Interspouse correlations for pulse

	ENG			NEW			OLD		
	n	r	SD	n	r	SD	n	r	SD
restp	67	.209	.121	110	.258	.093**	215	.098	.068
testp	67	.200	.122	109	.245	.094**	201	.064	.071
pratio	67	.160	.122	109	.136	.096	201	.083	.071
restp2	51	.068	.143	95	.182	.102	.194	.048	.072
testp2	51	.241	.139	94	.266	.101**	183	-.037	.074
pratio2	51	-.142	.141	92	.034	.105	.181	-.004	.075
bestrestp	67	.184	.122	111	.175	.094	219	.053	.068
besttestp	67	.282	.119	111	.259	.093**	211	.012	.069
bestpratio	67	.056	.124	112	.048	.095	211	-.038	.069

restp = resting pulse; testp = after exercise; pratio = ratio testp/restp

TABLE 4.12

Interspouse correlations for blood pressure

	ENG			NEW			OLD		
	n	r	SD	n	r	SD	n	r	SD
sysbp	67	.099	.123	111	.218	.093**	215	.258	.066****
diabp	67	-.116	.123	111	.127	.095	214	.043	.069
sysbp2	51	.170	.141	97	-.005	.103	193	.099	.072
diabp2	51	.253	.138	97	.093	.102	193	-.025	.072
bestsysbp	67	.213	.121	112	.045	.095	219	.135	.067*
bestdiabp	67	.151	.123	112	.112	.095	218	.005	.068

sysbp = systolic;   diabp = diastolic

IV. DISCUSSION

Physical characteristics were included in this survey for two reasons - to enquire into the nature of assortative marriage for height, and to gauge the effect of similarity of lifestyle, as reflected in physique.

In the case of height the findings were fairly clear-cut. There was quite strong ($r \approx .26$) assortative marriage for height. There was an intercorrelation with age and with social class, especially in older couples, suggesting a nutritional effect or the toll of degenerative bone or joint disease. However none of the other traits examined could account for the similarity for height. As this similarity was found in all three groups it was probably an important factor in mate selection. There is also the complex interaction of height and social class discussed by Wilson and Nias. It is likely that we are dealing with three overlapping factors here:

a) Strong selection for relative height, so that shorter men are rejected.

b) Directional selection for tallness, which is probably considered attractive in both sexes.

c) Normalising selection, so that gross stunting or gigantism are rejected.

In fact Mitton (1975) demonstrated strong normalising selection for height of men though not of women. One might expect assortative marriage (if it has any biological role at all) to minimise the occurence of pelvic disproportion. A small woman can not deliver a large baby through her small pelvis and until recent generations she would have died in the attempt. In that case the combination to be avoided would be tall men marrying small women. But it is unlikely that this has ever been biologically important. Pelvic disproportion has much more to do with nutrition of mother and fetus than with inherited stature.

Additionally, people are altering their own height by choice of shoes, though no tendency to mis-estimate height was found. It is possible that the very short or tall were self-conscious about this but Table 4.3 revealed no link between height and neuroticism. Wilson and Nias (1976) argued that it is the points of maximum difference in physique that are the most sexually arousing, and that the idea of beauty treatment is to emphasise these: lips, complexion, hair, etc. In that case one might wonder why women, rather than men, wear high-heeled shoes. Although the dictates of fashion must play a part high heels are such a long established custom that it must be women's consumer pressure that has influenced the fashion industry, not vice versa. High heels must have a large pay-off to the wearer, to compensate for the loss of height difference (not to mention their inherent awkwardness), and this pay-off is presumably in enhanced sexual attractiveness. A woman who found herself attracted to a shorter man could always discard her high heels in order to preserve the conventional fiction that the man be the dominant member of the pair. But if she chose a man at least 2" taller (as about 80% of this sample have) she could wear high heels and still not tower above her partner.

As height is the only trait in the study which is almost entirely under genetic control, we ought to consider the genetic consequences of assortative marriage, (Although there are still nutritional causes of stunting they are disappearing from Western populations. In that case the environment becomes effectively constant and the heritability rises towards 100%).

The weighted mean of the three correlations for "best" height was r = 0.266 $\pm$ 0.049. This is in the same range as the values listed by Roberts (1977). On the basis of previous work it has been calculated that assortative marriage contributes 17% of the variance of height. (Cavalli-Sforza and Bodmer 1971). That result assumes that assortative marriage for height has been constant for many generations so that there is now an equilibrium. Changes in heritability as the population approaches its environmental 'ceiling' might be one disturbing factor but in fact the figure of 17% was arrived at by assuming no environmental contribution. Harrison et al (1976) found homogamy for height only in classes IV and V in Oxfordshire and that would have been important if generally true, implying that changing British class-structure would lead to changes in assortative marriage. But in the present study if all those couples where either the man or the woman was in Class IV or V are lumped together, to get adequate numbers, the result is $r=0.174 \pm .132$ (n = 58). The bulk of the association is coming from the middle and upper classes. Probably therefore in Britain as a whole there is no difference among the classes for homogamy for height, and the 17% equilibrium figure is valid.

This increase in variance is counteracted by normalising selection as Milton (1975) reported. An attempt was made to confirm this finding. Simple correlations, such as 'his bestht v his totkids2' can only detect linear trends such as directional selection and results were close to zero in all three cohorts (Table 4.3). To test for normalising selection we have to split the sample into those above and below average height. Those below should have a positive, and those above a negative correlation between height and fertility. Although the correlations found (not tabulated) were mostly in the direction expected, there was only one result significant at the p = 0.05 level out of the 24 correlations calculated. Therefore we can not claim to have demonstrated normalising selection in the present sample.

Pomerat (1936) had found the remarkable correlation of 0.63 for height in infertile couples. If the best matched couples were infertile this would lead to strong divergent selection which seems highly unlikely. Unfortunately data from a modern comparable group is not yet to hand, as it is only in recent weeks that the Edinburgh Infertility Clinic has routinely measured the heights of couples. It will be possible to make a comparison by the end of 1980.

Other measurements of physique showed scant tendency towards assortative marriage. It had been hypothesized, for instance, that the correlations for weight might have been accounted for by correlations for skinfold and height; but here we have no correlation for weight to account for. The relative weights were therefore calculated, and still little correlation found. In all three of weight, relative weight and skinfold thickness there may be a transient correlation shortly after marriage but this can not be regarded as three independent lines of evidence all leading towards the same conclusion, as the three are highly correlated.

Unlike height, weight and skinfold thickness, pulse and blood pressure are not phenotypically obvious. It was clear at the outset that these traits could only indirectly measure life style. Consequently, the rather slight degree of assortative marriage found was no surprise, although it is clear that interests in sports etc. are an important factor in choosing a partner (Kreitman 1964). If the significant results in Table 4.12 had genuinely been reflecting assortative marriage for athleticism, one would have expected to see the highest correlations among engaged couples. It appears that relatively few married women practise any regular sport, and obviously once children are born it is very difficult for them to do so.

Although physical traits have proved uninformative as a measure of "healthy living" there are other fairly simple indicators. It will be shown in Chapter 6 that smoking is an example.

CHAPTER 5.

Psychological Traits.

I. Introduction

II. Methods

III. Results

IV. Discussion

CHAPTER FIVE

I. INTRODUCTION

It is among the psychological traits that changes during the course of a marriage might be particularly expected. Whilst both common sense and formal measurement tell us that these traits are important in selecting a partner, they are overshadowed to some extent by assortative marriage for age and social class. But in no other group of traits can incompatibility so quickly disrupt a marriage. There may be changes as partners adjust to each other, and to their altering roles as the marriage progresses. (Normal and abnormal family psychology has recently been reviewed by Dominian 1979 a-k). It is important to distinguish between similarity and compatibility. In some traits it may be that opposites attract as Winch (1955) has claimed, or that the degree of similarity that leads to optimum compatibility varies during different stages of marriage.

As well as personality traits, 'psychology' would include IQ, memory and perhaps personal habits such as smoking or drinking, as well as attitudes to these traits as they are perceived in the other partner. In this study IQ was only measured very indirectly, via educational attainments. As both education and smoking habits are closely linked to social class it is more convenient to discuss them together, and this will be done in Chapter 6.

The present chapter therefore deals only with the personality traits, neuroticism and extroversion, as measured in the third part of the interview.

II. METHODS

The short form of Eysenck's Personality Inventory (EPI) was used. (See appendix). This twelve-question inventory gives a reasonable measure of neuroticism and extroversion: a sample of 1600 men and women showed good clustering of the answers on scales for neuroticism and extroversion, with no intercorrelation of the two characteristics (Eysenck 1960). The EPI was developed from the Maudsley Personality Inventory (MPI), which had the drawbacks that some of the questions were too complicated for some people to understand, and that 'neuroticism' and 'extroversion' as measured by the MPI were correlated (Eysenck & Eysenck 1964). However even in the improved EPI there remains a correlation with age (extroversion: $r = -.241$, neuroticism: $r = -.163$), though not with sex.

A technical problem with this and similar psychological tests is that of 'response sets'. A response set is a tendency to respond in a stereotyped way which is more or less independent of the actual questions asked. For example, the 'acquiescence set' means the tendency to answer 'yes' to every question, though this has not been found to be a problem in practice. Another example is the 'desirability set' - the tendency to mark the answers which the respondent thinks the tester is looking for. Some versions of the EPI include a 'lie-scale' to check if this is happening. However this is only an important factor when the test is being used as a selection procedure, which is not the case here, and the version of EPI used here does not have a lie-scale.

For each person, the extroversion score was totalled as the number of 'yes' answers to questions 1, 4, 5, 8, 9 and 12, plus a half-point for each question-mark ringed. Likewise the neuroticism score was totalled from answers number 2, 3, 6, 7, 10 and 11. Thus for both traits the total score ranged from zero to six. (This scoring was done within the computer program DANTE, with the raw answers forming the input.) It was immediately apparent on looking at the completed questionnaires that the numbers of question-marks ringed by each partner were very similar. The tendency to mark the '?' is a response set rather than a personality trait (Eysenck 1962). However, having found strong evidence of assortative marriage in this unexpected source, it would be foolish to ignore it. The name QUERO was therefore coined to describe this trait/tendency by analogy with EXTRO and NEURO which were the names by which extroversion and neuroticism were identified to the computer.

Each person then rated his/her partner for neuroticism and extroversion, using the same EPI. It should be noted that

a) they were instructed to mark what they thought was true of the partner, not what they thought the partner would answer,

b) they were not aware of what the partner had answered.

These cross-ratings were termed NEUROX, EXTROX and QUEROX, thus 'his neurox' means the man's assessment of the woman's neuroticism; 'her querox' means the number of '?'s marked by the woman while assessing the man, and so on.

In measuring cross-assessments, it is possible that correlations may be exaggerated by a 'similarity set' - a tendency to answer for one's partner as for oneself, in the belief that close agreement shows the marriage in a good light, as harmonious. Another response set with indistinguishable results would be 'projection' - the tendency to ascribe one's own personality traits to the partner. This must be occurring to some extent: one man, when asked to assess his wife, instantly said, "Oh they're all exactly the same" without even glancing at the question sheet. However there are ways of gauging whether this is a serious problem and these will be demonstrated below.

III. RESULTS

Table 5.1 summarises the correlations found for psychological traits during the first year's visits. The first three lines show the strength of assortative marriage for neuroticism, extroversion and the 'Q' trait. Neuroticism was correlated in all three cohorts (.247 engaged, .330 new, .237 old). Extroversion was significantly correlated only in newly weds (.097, .290, .039). The 'Q' trait was strongly correlated in all three cohorts but again more so in newly weds (.557, .660, .549).

The calculation of 'his neuro v her neurox' measures how accurately the man's assessment of his own neuroticism matched the woman's assessment of him. Similarily 'her neuro v his neurox' gauges the agreement between the woman's self-assessment and the man's assessment of her. We could have called these 'subjective' and 'objective' ratings respectively; but 'objective' also carries connotations of increased accuracy through logic and detachment. In fact there was no reason to regard either form of assessment as inherently more reliable than the other. Nor should we consider them as but imperfect mirrors of some absolute

TABLE 5.1

Correlation for psychological traits measured at the first visit

	ENG (n=68)		NEW (n=113)		OLD (n=221)	
	r	SD	r	SD	r	SD
his neuro v her neuro	.247	.119*	.330	.090***	.237	.066***
his extro v her extro	.097	.123	.290	.091**	.039	.068
his quero v her quero	.557	.102****	.660	.071****	.549	.056****
his neuro v her neurox	.229	.120	.484	.083****	.326	.064****
his neuro v his neurox	.087	.123	.143	.094	.231	.066***
her neuro v his neurox	.450	.110****	.312	.090***	.300	.064****
her neuro v her neurox	.096	.123	.020	.095	-.066	.067
his neuro v his extro	.018	.123	-.038	.095	-.082	.067
her neuro v her extro	-.205	.120	-.139	.094	-.076	.067
her quero v her extro	-.009	.123	.014	.095	-.149	.067*
his quero v his extro	-.133	.122	.010	.095	-.017	.068
his neuro v his quero	-.126	.122	.105	.094	-.099	.067
her neuro v her quero	.024	.123	-.051	.095	-.066	.067
his extro v his extrox	.117	.122	.125	.094	.071	.067
his extro v her extrox	.559	.102****	.534	.080****	.412	.062****
her extro v her extrox	-.051	.123	-.007	.095	-.218	.066***
her extro v his extrox	.353	.115**	.427	.086****	.446	.060****

NEURO - 'how I rate my own neuroticism'. NEUROX - 'how I rate my partner's neuroticism'.
EXTRO - 'how I rate my own extroversion' EXTROX - 'how I rate my partner's extroversion'
QUERO - 'don't knows' when rating self QUEROX - 'don't knows' when rating partner.

psychological truth but as valid perceptions in their own right.

The cross-assessments turned out to be highly correlated, though in the latter case the engaged couples matched closer than the newly weds. Even better agreement was shown when the same calculations were done for extroversion ('his extro v her extrox' and 'her extro v his extrox' at the foot of Table 5.1). Thus, people's self-perceptions were well matched to their partner's perceptions of them.

Calculation of 'his neuro v his neurox' measures the agreement between the man's assessment of his own neuroticism and his assessment of the woman's neuroticism. A small positive value was to be expected, because we have already seen above that 'his neuro v her neuro' and 'her neuro v his neurox' were both correlated. But a large value might suggest that some other factor e.g. 'projection' or the 'similarity set' had spuriously inflated the correlation.

In fact the correlation was not significant in engaged and newly wed couples, but had a high ($r = .231 \pm .066$***) value in old men. The women's results (her neuro v her neurox) were never significant. But in the case of extroversion, the men (his extro v his extrox) showed nothing significant, whereas the women (her extro v her extrox) had a correlation only in the older group, but a negative one ($r = -.218 \pm .066$***). We can say at once that 'projection' and the 'similarity set' have had little influence on engaged and newly wed couples but may have been relevant to the older group. This group was therefore examined further by making correction for the small positive correlation predicted above. In the first case, 'his neuro v his neurox' corrected for 'her neuro' fell to $r = .158 \pm .067$*. In the second, 'her extro v her extrox' corrected for 'his extro' was $r = -.242 \pm .066$***. So the correlation

TABLE 5.2

Correlations for psychological traits measured at the second visit

	ENG (n=52)		NEW (n=98)		OLD (n=197)	
	r	SD	r	SD	r	SD
his neuro2 v her neuro2	.392	.130**	.228	.099*	.278	.069****
his extro2 v her extro2	-.020	.141	.291	.098**	-.159	.071*
his quero2 v her quero2	.787	.087****	.575	.083****	.668	.053****
his neuro2 v her neurox2	.036	.141	.343	.096***	.358	.067****
his neuro2 v his neurox2	.080	.141	.215	.100*	.139	.071
her neuro2 v his neurox2	.366	.132**	.226	.099*	.251	.069***
her neuro2 v her neurox2	.040	.141	.016	.102	-.078	.072
his neuro2 v his extro2	.019	.141	-.172	.101	-.115	.071
her neuro2 v her extro2	-.353	.132**	.007	.102	-.148	.071*
her quero2 v her extro2	.006	.141	-.011	.102	-.093	.071
his quero2 v his extro2	-.040	.141	.001	.102	-.049	.072
his neuro2 v his quero2	.051	.141	-.133	.101	-.117	.071
her neuro2 v her quero2	-.079	.141	-.205	.100*	-.073	.071
his extro2 v his extrox 2	.314	.134*	.260	.099**	-.011	.072
his extro2 v her extrox2	.650	.108****	.445	.091****	.378	.066****
her extro2 v her extrox2	-.126	.140	.135	.101	-.230	.070**
her extro2 v his extrox2	.407	.129**	.335	.096***	.430	.065****

for neuroticism remained significant but small, while that for extroversion increased its negative value. Projection/similarity set does not appear to be of practical importance.

Table 5.1 confirms that neuroticism and extroversion were not intercorrelated in this sample. The 'Q' trait also appeared to be an independent factor, though 'her quero v her extro' showed a borderline correlation in older couples. As Table 5.1 contains 51 correlations, and as 'his quero v his extro' was almost zero, it seems reasonable to dismiss this as a 'false-positive' correlation.

Table 5.2 presents exactly the same calculations as in Table 5.1, but based on the repeat visits one year later. Again the 'Q' trait showed the highest correlation and extroversion was matched only in the newly weds. (Indeed, there is a negative correlation in old couples.) The correlation for neuroticism changed little. Accuracy of assessment of partners remained high for both neuroticism and extroversion, and as before the three personality traits appeared to be independent.

Table 5.3 shows the repeatability of the personality test between the two visits. Neuroticism and extroversion showed a fairly uniform repeatability of 70%, which is comparable to the Eysencks' own finding of about 85% after a four week interval (Eysenck & Eysenck 1964). But the 'Q' trait showed much lower repeatability, suggesting something much more ephemeral than a true personality trait.

Table 5.4 shows the mean and standard deviation of the various scores. All the measures of neuroticism and extroversion had means close to 3, which is in the middle of the range of possible scores. There is nothing to suggest that scores were increasing or decreasing with length of marriage. The 'Q' trait had a skew distribution, with mode zero and median about 0.5.

TABLE 5.3

Repeatability of the psychological ratings
i.e. comparing the same trait for the first and second visit

	ENG (n=52)		NEW (n=97)		OLD (n=196)	
	r	SD	r	SD	r	SD
his neuro v his neuro2	.705	.100****	.713	.072****	.725	.049****
her neuro v her neuro2	.521	.121****	.499	.088****	.725	.049****
his neurox v his neurox2	.730	.097****	.637	.078****	.755	.047****
her neurox v her neurox2	.700	.101****	.715	.071****	.659	.054****
his extro v his extro2	.840	.077****	.579	.083****	.700	.051****
her extro v her extro2	.575	.116.****	.598	.082****	.666	.054****
his extro v his extrox2	.664	.106****	.667	.076****	.691	.052****
her extrox v her extrox2	.602	.113****	.773	.064****	.700	.050****
his quero v his quero2	.219	.138	.277	.098*	.397	.066****
her quero v her quero2	.351	.132**	.302	.097*	.514	.062****
his querox v his querox2	.359	.132**	.348	.095***	.426	.065****
her querox v her querox2	.552	.118****	.300	.097**	.329	.068****

Table 3.7

Mean and SD of neuro, extro and 'Q' scores

	ENG		NEW		OLD	
	mean	SD	mean	SD	mean	SD
his neuro	3.28	1.79	3.33	1.80	2.94	1.84
her neuro	2.94	1.69	3.80	1.60	2.90	1.77
his neurox	3.26	1.77	4.11	1.51	3.42	1.75
her neurox	2.41	1.85	2.87	1.80	2.48	1.73
his neuro2	2.91	1.99	3.17	1.91	2.89	1.91
her neuro2	2.95	1.50	3.61	1.57	2.80	1.88
his neurox2	3.38	1.87	4.17	1.66	3.52	1.81
her neurox 2	2.12	1.66	2.87	1.85	2.44	1.76
his extro	3.99	1.38	3.76	1.52	3.62	1.53
her extro	3.51	1.63	3.53	1.59	3.55	1.60
his extrox	3.90	1.52	3.85	1.34	3.57	1.55
her extrox	4.31	1.55	3.82	1.65	3.67	1.63
his extro2	3.92	1.66	3.83	1.59	3.78	1.41
her extro2	3.55	1.65	3.57	1.53	3.52	1.64
his extrox2	3.64	1.44	3.64	1.49	3.56	1.66
her extrox2	4.26	1.46	3.52	1.74	3.43	1.61
his quero	0.44	0.66	0.66	1.17	0.59	0.98
her quero	0.74	1.27	0.58	1.00	0.40	0.88
his querox	0.63	0.84	0.68	1.09	0.44	0.85
her querox	0.53	0.85	0.52	1.01	0.31	0.83
his quero2	0.67	1.26	0.55	0.77	0.55	0.98
her quero2	0.73	1.27	0.58	1.05	0.45	0.97
his querox2	0.71	0.96	0.68	1.00	0.30	0.63
her querox2	0.62	1.16	0.48	0.94	0.23	0.74

Tables 5.5, 5.6 and 5.7 summarise the intercorrelations that were found between the personality factors and the other traits studied. These are mostly nonsignificant. The negative correlation with age predicted by Eysenck was only significant for the neuroticism of engaged women and old men. Otherwise the only finding of note was in the newly wed men: the better educated were more introvert ($r = -.322 \pm .090$***). The 'Q' trait was correlated with education and social class in older men, perhaps with religion and height in others. That is to say, the better educated, higher social class people (who were more likely to practise a religion, and in old couples likelier to be tall - Table 4.3) were more likely to ring the question mark.

IV. DISCUSSION

The set of possible cross-correlations forms an infinite regress: "How I see myself", "How I see my partner", "How I believe my partner sees me", "How I believe my partner believes I see him/her", and so on. Another dimension lies in the subjunctive: "Myself as I would like to be", "My partner as I would like him/her to be" and so ad infinitum. Only the first four levels of introspection seem of practical import. In fact rather few investigators have penetrated beyond the first level, and it appears that only one study (Drewery and Rae 1969) has reached the third. The methodology is therefore less well tested for cross-assessments. Nevertheless it is of some importance because here we are dealing with a group of traits where assortative marriage contributes to the success of the partnership, as Tharp (1963) and Luckey (1960) have confirmed. There were not enough split-ups in the present study to show this however. But it could be demonstrated that spurious factors such as response sets did not seem to be greatly influencing the results. There remains the problem of repeatability. In this and other psychological traits (especially IQ) results may be altered simply by practice at taking the test.

TABLE 5.5

Correlation of neuroticism versus other traits

	ENG				NEW				OLD			
	M		F		M		F		M		F	
	r	SD	r	SD	r	SD	r	SD	r	SD	r	SD
age now	.097	.123	-.277	.118*	-.111	.094	-.176	.093	-.236	.066***	.027	.068
bestjob	.103	.122	.131	.122	-.012	.095	-.021	.095	.147	.067*	.067	.067
smoking	-.110	.122	.099	.122	-.025	.095	-.065	.095	.154	.067*	.079	.067
religion	-.077	.123	-.156	.122	-.036	.096	-.133	.095	-.201	.066**	-.016	.068
education	-.085	.123	-.059	.123	-.063	.095	.043	.095	-.125	.067	-.063	.067
bestht	-.046	.123	.121	.122	-.066	.095	.076	.095	-.025	.068	-.083	.067
bestwt	-.161	.121	.111	.122	.054	.095	.040	.095	-.065	.067	.035	.068

'bestjob' is social class as defined in Chapter 6.
'bestht' and 'bestwt' are best estimates of height and weight as defined in Chapter 4.
Column 'M' is for male values - his neuro v his age now, etc. and Column F for female values.

TABLE 5.6

Correlation of extroversion versus other traits

	ENG				NEW				OLD			
	M		F		M		F		M		F	
	r	SD	r	SD	r	SD	r	SD	r	SD	r	SD
age now	-.170	.121	.148	.122	-.134	.094	-.019	.095	-.087	.067	.044	.068
bestjob	.254	.119*	.007	.123	.181	.093	.108	.094	.107	.067	-.024	.068
smoking	.224	.120	-.151	.122	.008	.095	.195	.093*	.127	.067	.039	.068
religion	-.165	.121	.173	.121	.052	.096	.088	.095	.011	.068	.105	.067
education	-.114	.122	-.034	.123	-.322	.090***	-.134	.094	-.122	.067	.010	.068
bestht	-.080	.123	.071	.123	-.068	.095	-.184	.093*	.025	.068	-.085	.067
bestwt	.013	.123	.274	.118*	.009	.095	-.130	.094	.014	.068	.159	.067*

TABLE 5.7

Correlation of 'Q' versus other traits

	ENG				NEW				OLD			
	M		F		M		F		M		F	
	r	SD	r	SD	r	SD	r	SD	r	SD	r	SD
age now	-.014	.123	.197	.121	.089	.095	.157	.094	-.110	.067	-.091	.067
bestjob	-.144	.122	-.230	.120	-.164	.094	-.003	.095	-.221	.066***	-.098	.067
smoking	.049	.123	-.089	.123	-.116	.094	-.035	.095	-.101	.067	-.048	.067
religion	.284	.118*	-.011	.123	.164	.094	.198	.094*	-.071	.067	-.069	.067
education	.199	.121	.162	.121	.039	.095	-.064	.095	.237	.066***	.063	.067*
bestht	-.003	.123	-.043	.123	.081	.095	.010	.095	.056	.067	.200	.066**
bestwt	-.111	.122	-.096	.123	.115	.094	.070	.095	.041	.068	-.075	.067

The Eysencks (1964) suggested that neuroticism reflects an inherited lability of the autonomic nervous system, whilst extroversion reflects the degree of stimulation of the central nervous system induced by the reticular formation. The precise neuro-anatomical wellsprings of these traits fortunately do not concern us. The chief question is whether the traits are deeply ingrained and immutable, or only change slowly with time, or whether they are fairly plastic. If they are fixed then the changes seen over the year merely reflect the unreliability of the test. It would not matter whether this fixing was due to genes, to early childhood experience, or to maturation in adolescence, so long as the process was complete in people old enough to participate in the study.

Provided that enough time had elapsed for couples to forget the questions completely, the repeatability of the test should not depend on the length of intervening time, if the trait is fixed. But if the trait alters, the longer the interval the lower the repeatability. A 70% repeatability over one year, set against Eysenck's finding of 85% over a month, suggests that neuroticism and extroversion were slowly but genuinely changing. Then engaged and newly wed couples would be expected to show the greatest change and hence the lowest repeatability. There is some evidence of this among the women: 'her neuro v her neuro2' being only 40-50% in new and engaged couples. Certainly the couples had forgotten the questions - they would usually say "Are these the same questions as last year?". And they would often have forgotten that after rating themselves they would be asked to rate their partner.

Kreitman (1964) found that in normal couples, there was initially a high degree of correlation for personality factors but that this declined slightly with time. In the present study, the correlation for neuroticism remained fairly steady, in the range $r = 0.25 - 0.3$. But the correlation for extroversion was high only in newly weds. Is it possible that couples were actually diverging, so that after six years the similarity vanished? Note that in older couples, the correlation actually became significantly negative by the second visit (Table 5.2). But we ought to have seen a higher correlation among those engaged couples who subsequently married. Perhaps the correlation was a peculiarity of this particular newly wed group. Previous workers have found little or no correlation for intro/extroversion (Vandenberg 1972). Whilst personality factors may have biased ascertainment in a way that would affect the strength of correlation found (e.g. if the more extrovert people were likelier to take part, and concordantly extrovert couples especially so) this bias should have applied equally to all three cohorts.

The principal difference between older couples and the others is that most (85%) older couples had children. They had therefore embarked upon the second stage of marriage, where the woman's interests are centred on home and family whilst the man's interests lie in his career. It therefore seemed worthwhile to attempt a separate analysis of the transition between the first and second stage, i.e. to focus on those couples whose first child was born between the two visits. However, there were only 22 of these and the standard deviations of the measurements were correspondingly high. Therefore it was not possible to demonstrate psychological differences between them and the main sample.

CHAPTER SIX.

Social Traits

I. Introduction

II. Methods

- a) Social Class
- b) Education
- c) Smoking
- d) Religion

III. Results

IV. Discussion

CHAPTER SIX

I. INTRODUCTION

This chapter is concerned with social class, education, smoking habits, and religion. In all of these a high correlation between spouses was to be expected, but this might be partly spurious as all the traits are closely interlinked. An attempt has therefore been made to unravel the intercorrelation between the various traits.

II. METHODS

a) SOCIAL CLASS

The 1971 Registrar General's Classfication offered a number of ways in which a person's status might be described. The traditional method is in terms of Classes. These are I (professional), II (managerial), IIIN (Clerical), IIIM (skilled manual), IV (semi-skilled) and V (unskilled). There is also the 17-part division into Socio-Economic Groups (farm and forestry, mining, transport, managerial, etc.). Although the Groups are in several ways more informative than the Classes, it is the Classes which seemed best to serve the present analysis. For instance, using the Groups a driver for the Coal Board would be graded differently from a bus driver; yet both are in Social Class IIIM. It is the similarity, not the difference, which seemed most relevant here. There was also the advantage that Social Classes are much easier to work with when computing correlations.

However, there are a number of ways in which the Classification is deficient or outdated. No doubt many of these deficiencies will be corrected when the next Classification is issued, during the next National Census, but that will be too late for the present purpose. For instance, students and servicemen are unclassified, perhaps reflecting society's ambivalence as to their true worth. Other occupations are lumped, which really require splitting, e.g. managers, nurses, accountants. In modifying the Classification, the guiding principle is that of comparability. If the person was not in his present job, what other sort of work might he be in, given his training and experience? The main changes between the Registrar General's and the present Classification are listed below.

1) Students: University undergraduates have been placed in Class II unless their course clearly led to a Class I job (doctor, dentist, vet, lawyer, chartered accountant, architect). Postgraduates went

into Class I unless they were clearly training for a Class II job (teacher, musician, commerce). Although, for example, "historian" is Class I, very few students doing MA History ever become historians - they become teachers or junior managers. Those doing Ph.D.-level history however would probably be aiming for university lecturing or research posts. But schoolchildren have remained unclassified.

2) Servicemen: General Staff, ship's captains and pilot officers went into Class I, other commissioned officers into Class II. Other ranks went into IIIM. Although they would have had very variable occupations and specialist skills they are to be considered, quite literally, uniformed.

3) Police and Fire: Chiefs went into Class I, others down to police inspector or station officer went into Class II. Everyone else went into Class IIIM except detective constables and sergeants, who went into IIIN.

4) Nurses remained in II, but unqualified auxiliaries went into IIIM.

5) Accountants went into II but Chartered Accountants into I.

6) Most managers remained in II. However there were a few entrepreneurs who dealt with colossal sums of money (ie would regard a transaction of less than £1m as not worth their while) and these have been placed in I. Assistant managers were demoted into IIIN.

7) Church ministers were demoted to II, where they seemed to fit better among teachers and social workers. Senior churchmen remained in I.

In making calculations, the question arose as to how Classes IIIN and IIIM should be distinguished. Class I is naturally represented as the number '1', II as '2' and IIIN as '3', but should IIIM be 3, or 4, or somewhere in between? In fact all three solutions were tried during the course of the survey. Initially IIIM was represented as '3' with IV as '4' and V as '5'. The reason for this was that the difference between IIIN and IIIM lies not so much in skill but in traditional sex roles. Also, at that early stage the numbers collected were small and it seemed best to lump categories. However, the gap between IIIN and IIIM is in many ways an important social divide. The analysis was therefore repeated with IIIM equivalent to 4, IV to 5, and V to 6. The correlations of, for example, smoking with social class, were greatly increased when this was done. But this seemed to go too far to the opposite extreme. It implied that the difference between an electrician (IIIM) and a clerk (IIIN) was as great as that between a clerk and a teacher; or the difference between a postman (IV) and a nurse (II) was like that between a typist (IIIN) and a lawyer (I). Thereafter a compromise was adopted: IIIM was scored as 3.5, IV as 4, and V as 5. All the results in the study have been based on those figures, but obviously one could haggle for different values which would alter the results. There remains the problem that no manual occupation, however skilled, can ever be considered more prestigious than the least skilled clerical job. This is not a defect of the Classification, but an anachronism of British society in general.

To maintain flexibility, the occupations in the data file were not encoded as numbers, but by letters K to Q, as shown below.

Letter	Mnemonic	Class	Number
Q	'query' - not known	-	-950
P	'professional'	I	1
O	'office' - managerial	II	2
N	'nonmanual' - clerical	IIIN	3
M	'manual' - skilled	IIIM	3.5
L	'labourer'- semi-skilled	IV	4
K	'kraftless' - unskilled	V	5

(Large negative numbers indicate unknowns or missing data; of course they did not enter into the calculations).

This meant that when changes were made in the evaluation of Classes it was only necessary to alter a single line of program instead of three bulky files.

b) EDUCATION

Educational levels were graded as follows:

5 - Postgraduate: i.e. having or studying for a higher degree such as M.Litt. or Ph.D, but not including postgraduate diplomas nor "second first" degrees such as B.D. following M.A.

4 - Undergraduate: having or studying for a university first degree.

3 - having Highers, A-levels, or the equivalent.

2 - having O-levels or the equivalent.

1 - no O-levels

0 - evidence of educational retardation (but making allowance for physical, sensory or linguistic handicap) - i.e. native English speakers unable to read, or to attend normal school.

Here, as in the classification of occupations, it was assumed that students taking University degrees would be successful. In fact, no student failed during the course of the survey.

Additionally, those educated before O-levels were introduced were upgraded to 2 if they stayed on at school after minimum leaving age, and to 3 if they stayed on after 16. In spite of these corrections there remained many older people in the survey who could only be graded 1 although they were undoubtedly intelligent. This could not be helped by regrading. The educational system of those days was inherently unfair and wasteful of talent, and the present grading system had to reflect that. We must bear that in mind when interpreting results.

c) SMOKING.

Smoking habits were coded as the number of cigarettes smoked per day, to the nearest whole number. However those smoking 3 or 4 per week were encoded as 0.5, those smoking less down to

4/month as 0.2. Anything less was coded as zero. A cut-off line had to be imposed because some people were very very light smokers (e.g. one cigar per Christmas and birthday) and for practical purposes counted as nonsmokers. The lightest smokers also found it hard to gauge their intake accurately because they were not regularly buying tobacco. The choice of cut-off line did not affect the correlation coefficient because among so many readings of 20 and 10 the effect of interchanging a few values between 0 and 0.5 was trivial. However it could have affected the concordance when smoking was treated as an all-or-none phenomenon. Therefore concordances were subsequently calculated over a range of cut-offs, to confirm that the original choice was justified.

There was also the problem that smokers often lie to doctors about their consumption (Kozlowski et al 1980). However, the strongest spur to deceit is in the anti-smoking clinic or GP's surgery, where they are anxious to avoid unfavourable medical comment. No such implied criticism occured during the present study, and the smoker's spouse made an extra witness.

One gram of tobacco is equivalent to one cigarette, so 1 oz. (28gm) of pipe tobacco/week was coded as 4 cigarettes/day. Cigars were treated as equivalent to cigarettes. One could argue for a differential in either direction. On the one hand, they are considered less of a health hazard than cigarettes, as less is inhaled. On the other, cigars are more expensive and perhaps more pungent to the nonsmoker.

d) RELIGION

Data on religion was encoded as his and her CHURCH. The all-or-none classification (1 = practising some religion, 0 = not practising) was adopted for quick testing purposes on the computer. This was adequate for checking, say, the intercorrelation of religion with social class, personality traits, etc. with the information on the actual religion practised being plotted out by hand.

TABLE 6.1

Assortative Marriage for Social Factors

	ENG n	ENG r	ENG SD	NEW n	NEW r	NEW SD	OLD n	OLD r	OLD SD
his bestjob v her bestjob	68	.298	.117*	113	.553	.079****	222	.488	.059****
his bestjob v his patjob	67	.445	.111****	112	.477	.084****	220	.413	.062****
his bestjob v her patjob	68	.437	.111****	113	.416	.086****	218	.467	.060****
her bestjob v her patjob	68	.220	.120	113	.399	.087****	218	.294	.065****
her bestjob v his patjob	67	.084	.124	112	.331	.090****	220	.388	.062****
his job v her job	66	.321	.118**	84	.397	.101****	105	.476	.087****
his job v his patjob	67	.314	.118**	105	.456	.088****	215	.413	.062****
his job v her patjob	68	.350	.115**	106	.384	.091****	213	.469	.061****
his job v his job2	52	.875	.068****	91	.951	.033****	187	.983	.013****
her job v her patjob	66	.240	.121*	86	.428	.099****	107	377	.090****
her job v his patjob	65	.145	.125	85	.207	.107	108	.494	.084****
her job v her job2	50	.967	.036****	68	.926	.046****	79	.913	.046****
his job2 v her job2	50	.303	.138*	75	.578	.095****	89	.482	.094****
his patjob v her patjob	67	.310	.118**	112	.327	.090***	218	.299	.065****
his school v her school	68	.478	.108***	113	.596	.076****	222	.640	.052****
his church v her church	68	.515	.106****	113	.415	.087****	222	.551	.056****
his church2 v her church2	52	.575	.116****	99	.422	.092****	198	.562	.059****
his smoking v her smoking	68	.431	.111****	113	.311	.090***	222	.314	.064****
his smoking2 v her smoking2	52	.523	.121****	99	.418	.092****	198	.255	.069****
his smoking2 v his smoking	52	.896	.063****	99	.780	.064****	198	.914	.029****
her smokin2 v her smoking	52	.946	.046****	99	.880	.048****	198	.917	.029****
smokers only: his v her smoking	not computable - only 9 couples			38	.004	.167	55	.333	.130*

III. RESULTS

Table 6.1 summarises the interspouse correlations found for occupation, education, religion and smoking habits. All of these traits showed clear evidence of assortative marriage with r = 0.4, and the value was similar in each cohort. Let us begin by looking more closely at measurements of social class.

The correlation for best job was higher than for father's job in new and old couples, though not in the engaged. This suggested that it was no longer true to say that a bride's social status was defined by that of her father: her own occupation was at least as good an indicator. Note however that in the engaged couples, his bestjob v her patjob was actually higher than his bestjob v her bestjob. This was investigated further by correcting bestjob for patjob. In every cohort the correlation for patjob became non-significant whereas that for bestjob remained high (Table 6.2). This suggested that it was the association for bestjob that induced that for patjob, not vice versa, i.e. couples are assorting according to their own abilities and not their family backgrounds. Results were similar when patjob was corrected for education.

Tables 6.2 - 6.5 show cross tabulations, i.e. who married who in each social class. For this trait we also know the composition of almost the entire marrying population from registry office notices. The cross tabulations for their social class is given in Table 6.5; their correlation was n=412 $r = 0.548 \pm 0.041$, a result very similar to that for the responders. The total and percent columns allow one to see the social class profile of each cohort and sex. The bias in response rate towards classes I and II is clear; but only in the engaged couples is the effect gross. The bias means that if a trait displayed a markedly different degree of assortative marriage in different social strata, the strength of correlation found in this study would not be typical of the population at large. Note

TABLE 6.2

Cross tabulation for social class : ENGAGED

		HER BESTJOB							
		I	II	IIIN	IIIM	IV	V	TOTAL	%
HIS BEST JOB	I	2	12	2				16	24
	II	1	22	6				29	43
	IIIN	1	9	3				13	19
	IIIM		5	1	3			9	13
	IV			1				1	1
	V							0	0
	TOTAL	4	48	13	3	0	0	68	
	%	6	71	19	4	0	0		100

TABLE 6.3

Cross tabulation for social class : NEW

		HER BESTJOB							
		I	II	IIIN	IIIM	IV	V	TOTAL	%
HIS	I	5	12	3	1			21	19
BEST	II	2	16	16				34	30
JOB	IIIN		2	7	2	1		12	11
	IIIM		12	16	4	3	1	36	31
	IV			2	1	1	2	6	5
	V			2		1	1	4	4
	TOTAL:	7	42	46	8	6	4	113	
	%	6	37	41	7	5	4		100

TABLE 6.4

Cross tabulation for social class : OLD

HER BESTJOB

		I	II	IIIN	IIIM	IV	V	TOTAL	%
HIS BEST JOB	I	5	19	7				31	14
	II	1	32	32	6	2		73	33
	IIIN		6	15	2	5		28	13
	IIIM		16	29	10	11	4	70	31
	IV		1	7	1	5	2	16	7
	V			3	1			4	2
	TOTAL:	6	74	93	20	23	6	222	
	%	3	33	42	9	11	3		100

TABLE 6.5

Cross tabulation for social class - all newly weds (responders and non responders)

		HER JOB AT MARRIAGE							
		I	II	IIIN	IIIM	IV	V	TOTAL	%
HIS JOB	I	10	28	9		1		58	12
	II	7	49	39	4	6		105	25
	IIIN		11	22	4	2		39	9
	IIIM	1	22	73	18	38	1	153	37
	IV		4	28	11	17		60	15
	V			3	1	3		7	2
	TOTAL	18	114	174	38	67	1	412	
	%	4	28	43	8	17	0		100

MEAN and SD of social classes: men 2.88 ± 1.01; women 2.85 ± 0.79

that in setting Table 6.5 against 6.2 to 6.4 we are not strictly comparing like with like. In the newly-wed non-responders only occupation at time of marriage is known. Women who were not working (usually marrying housewifes) therefore could not be coded. The means and SDs of the occupational and educational levels of the three cohorts are given in Table 6.6. Again, the engaged couples are the least typical, with mean Social Class of 2.18, compared to a population value of 2.87.

By comparison, those old couples who announced an engagement (n=56) had mean Social Class of 2.01 $\pm$ 0.90 (men) and 2.50 $\pm$ 0.62 (women). Their parents' social classes were 2.62 $\pm$ 0.90 (men's fathers) and 2.54 $\pm$ 0.82 (women's fathers). These classes are higher than those for the old cohort in general, confirming that getting engaged is a middle class phenomena, but it seems to reflect the couples' own status as much as that of their parents.

The couples seen at the Infertility Clinic in Edinburgh also had a high correlation for social class: n = 77 r = .620 $\pm$ .091**** (The missing 23 were only secondarily infertile, e.g. they had had children with a previous partner, and therefore the women could only be categorised as housewives.)

Tables 6.7 - 6.9 show cross tabulations for education. The strong diagonal layout illustrates the correlation (.478 engaged, .596 new, .640 old) which is even higher than that for social class. The average participant reached O-level standard in the case of old couples, A or H level in the engaged, with newly-weds halfway between. Information on non-responders was not available. (Attempts to infer education from job, e.g. how many were grad-udates, proved inaccurate.) Yet when correction for social class was made (Table 6.25) it appeared that the correlation of social class was merely a reflection of association for education. This surprising result will be discussed later.

TABLE 6.6

Social Class and Education distributions : mean $\pm$SD

	ENG		NEW		OLD	
his bestjob	2.18	0.87	2.61	1.09	2.82	0.80
her bestjob	2.18	0.55	2.66	0.85	2.66	1.00
his patjob	2.43	0.93	2.76	1.06	3.07	0.94
her patjob	2.18	0.89	2.69	1.15	3.07	0.91
his education	3.24	1.02	2.36	1.34	2.05	1.23
her education	3.09	0.84	2.46	1.17	2.02	1.11

TABLE 6.7

Cross-tabulation for education : ENGAGED

		HER SCHOOL 0	1	2	3	4	5	TOTAL	%
HIS SCHOOL	0							0	0
	1		1	2	1	1		5	7
	2		3	1	3	2		9	13
	3			4	17	3		24	36
	4			1	11	13		25	37
	5				3	1	1	5	7
TOTAL		0	4	8	35	20	1	68	
	%	0	6	12	51	30	1		100

Where educational levels are: O: could not attend normal school
1: normal school, no 0-levels
2: O-levels
3: H/A - levels
4: university degree
5: postgraduate degree

TABLE 6.8

Cross tabulations for Education : NEW

		HER SCHOOL							
		0	1	2	3	4	5	TOTAL	%
	0		1					1	1
HIS	1		24	9	9	3		46	40
SCHOOL	2		3	5	6			14	12
	3		3	7	10	2	1	23	20
	4		2	1	7	14	1	25	22
	5				2	2	1	5	5
	TOTAL	0	33	22	34	21	3	113	
	%	0	29	19	30	19	3		100

TABLE 6.9

Cross tabulations for Education : OLD

HER SCHOOL

		0	1	2	3	4	5	TOTAL	%
	0		2					2	1
	1		74	20	12	1		107	48
HIS	2		15	6	13	1		35	16
SCHOOL	3		11	11	14	2	1	39	18
	4		1	2	14	16	1	34	15
	5			1	2		2	5	2
	TOTAL	0	103	40	55	20	4	222	
	%	0	46	18	25	9	2		100

Table 6.1 shows that correlations for religious practice were high (.515 engaged, .415 new and .551 old), but more informative are Tables 6.10 - 6.12 which detail the religions practised, the changes over the year, and the concordances. Not surprisingly, the Church of Scotland (CS) was the most popular among those who practised a religion. CS included Presbyterian and Free Kirk just as CE included Church of England, Episcopalian and Anglican. RC was Roman Catholic. Category 'X' (Christian) appears separately because several people said that they considered themselves practising Christianity in a broad sense but would not wish to be placed in any particular denomination. 'Solo' was the term coined to describe those who said they practised their own particular brand of religion. The categories are therefore exclusive, although the theological differences between most of them are slight.

42% of engaged, 18% of newly wed and 29% of older men practised a religion and among the women it was 55%, 33% and 46%. The women were therefore noticeably more religious, and religious man v nonreligious woman partnerships particularly scarce.

The overall correlations and concordances showed almost no change over the year, although the individual changes listed beneath the tables showed a slight tendency to converge. This was most noticeable in the engaged, where two people converted to Catholicism in order to marry their partner. In the newly-weds there are surprisingly few concordantly Church of Scotland marriages. It was possible that these were selectively lost by giving banns in church. (Before 1978 only Church of Scotland marriages gave banns in church and were lost to ascertainment - all other denominations had to give notice at the registrar's office). Yet of those new couples ascertained by wedding notice alone (who therefore must have married in church), in 2 cases the wife alone practised any religion and in the other 5 neither did. Therefore, Church marriage was not related to active religious practice.

TABLE 6.10

Cross tabulation for Religion : ENGAGED

		CS	CE	RC	X	SOLO	NIL	OTHERS	TOTAL
	CS	9	1	3			2	1	16
	CE	1							1
	RC	1					1		2
MEN	X				4		1		5
	SOLO					1			1
	NIL	11			1	1	29		42
	OTHERS							1	1
	TOTAL	23	1	3	5	2	32	2	68

includes CS2 v CS2, CS2 v CS1, CS2 v CS1, O v CS2
Others: E10 CB v CB, E54 CS v Bapt.
Changes: E45 CS v RC -> RC v RC; E20 CS v CE -> CE v CE
E25 X v O -> X v X; E9 RC v O -> RC v RC

Concordance:	M-F O-O	M-F O-1	M-F 1-O	M-F 1-1	r
1st year	27(40%)	12(18%)	5(7%)	24(35%)	0.515
2nd year	19(37%)	10(19%)	2(4%)	21(41%)	0.575

KEY to Tables 6.10 - 6.12: CS: Church of Scotland, Presbyterian or Free Church; CE: Church of England, Episcopalian or Anglican; RC: Roman Catholic; X: Christian, no denomination; SOLO: practices own religion; OTHERS:- CB: Christian Brethren; Ortho: Greek Orthodox; CX: Church of Christ; Islam: Shiai Islam; Bapt: Baptist; Meth: Methodist. CS2: involved in ministry (e.g. above - one marriage with both involved in ministry, two with men involved, wives actively practise CS; one with non-practising man, wife involved in ministry.)

Changes: e.g. E45 CS v RC -> RC v RC means man in couple No. 45 converted from CS to RC during the year.

TABLE 6.11

Cross tabulation for Religion : NEW

WOMEN

MEN	CS	CE	RC	X	SOLO	NIL	OTHERS	TOTAL
CS	3							3
CE	1	3						4
RC			1			2		3
X				3		1		4
SOLO				2		2		4
NIL	11	4	5	2	1	69	1	93
OTHERS	1					1		2
TOTAL	16	7	6	7	1	75	1	113

No CS2 marriages

Others: N37 Ortho v CS N85 O v CX N106 ISLAM v O

Changes:
N19 O v O -> O v X
N45 O v O -> O v CS
N55 SOLO v O -> O v X
N85 O v CX - > O v O
N104 O v CS -> O v O
N 35 O v O -> X v X
N47 O v RC -> O v O
N61 O v O -> O v X
N93 O vO -> O v CE

	M-F	M-F	M-F	M-F	
Concordance	0-0	0-1	1-0	1-1	r
1st year	69(61%)	24(21%)	6(6%)	14(12%)	0.415
2nd year	61(61%)	22(22%)	3(3%)	13(13%)	0.422

TABLE 6.12

Cross tabulations for Religion : OLD

		WOMEN							
		CS	CE	RC	X	SOLO	NIL	OTHERS	TOTAL
MEN	CS	28		1	1		3		33
	CE	1	3						4
	RC	2		4			3		9
	X	1			8				9
	SOLO	2				1	1		4
	NIL	29	2	7	4	1	113	2	158
	OTHERS	1						4	5
	TOTAL	64	5	12	13	2	120	6	222

Includes two CS2 v CS1

Others: O37 Quaker v Quaker O43 Meth v Meth
O111 and O112 Bapt v Bapt O121 Copt v CS O125 and O158 O v Spiritualists

Changes: O11 O v CS -> O v O O14 O v O -> RC v O O40 RC v O -> RC v RC
O138 O v CS -> O v O O148 O v SOLO -> O v O O216 O v O -> O v CS

	M-F	M-F	M-F	M-F	
Concordance	0-0	0-1	1-0	1-1	r
1st year	113(51%)	45(20%)	7(3%)	57(26%)	0.551
2nd year	101(51%)	41(21%)	5(3%)	51(25%)	0.562

Because church-going seemed an inappropriate criterion for strength of religious practice, no questions were asked about this. However we can pick out certain people as particularly active in the church because they were ministers, kirk elders, etc. They are the ones marked CS2 below the tables, implying that if the average active member of the Church of Scotland is marked as 1, these people would merit marking 2. In fact, there were not enough of them to justify embarking on any special calculations. However, the engaged couples clearly emerged as the most religious, with five such people, including one 'assortative' CS2 v CS2 partnership. Table 6.24 shows the correlation for religion after social class had been accounted for. One might have expected a marked reduction but in fact religion, though strongly linked to social class, assorted independently.

However this finding could not be accepted at face value, because the correction formula is rather sensitive to the Normality (or lack of it) of the data; and classification of religion was all-or-none. Confirmation was therefore sought from another method.

The 'select' routine was used to subdivide each cohort by each social factor in turn, to demonstrate the correlation for religion that remained in each case.

('select only his bestjob < 3.3
split his church v her church 0.5 0.5
select only his bestjob > 3.3 ' and so on)

In fact the correlations (not tabulated) always remained significant, except in one small subgroup: newly weds where the wife was in social class IIIM, IV or V ($n=17$, $r = 0.436 \pm 0.232$). Therefore, the other factors do not account for the similarity for religion.

Conversely, the sample was then subdivided by religious practice, before evaluating the remaining correlation in each social factor

('select only his church = 1
correl his bestjob v her bestjob and so on)

Results appear in Table 6.25. Education is still highly correlated, and so is smoking in all but two subgroups. Father's occupation is never significant among religious men, but clearly father's occupation can influence one's own religion but not vice versa. The correlation for bestjob becomes borderline in subgroups of the engaged. In summary, it appears that AM for religion is independent of the other social factors.

Turning now to smoking habits, these showed a high correlation in all three cohorts, the mean correlation being $r = 0.332 \pm 0.047$. Table 6.13 shows the smoking habits of the various cohorts and sexes, the main finding being that much fewer engaged people smoked. From this table we can also see that if people had married at random with respect to smoking habits we would have expected both partners to be smokers in 5% engaged, 25% new and 17% old couples. The observed proportions were 12% engaged, 32% new and 24% old. The correlations a year later were similar (Table 6.1). Note that the repeatability (his smoking v his smoking2 and her smoking v her smoking2) was very high, about 90%. This implied that any change in the correlation over the year was unlikely to be due to true convergence. It simply reflected which couples were revisited earliest.

It was mentioned earlier that a cut-off level had to be imposed below which people were considered nonsmokers. Perhaps the level chosen was arbitrary. Table 6.17 was therefore compiled to show how the concordance might vary when differing levels were set: 0.1, 0.3, 0.9, 1.5, 4.9 and 19.5 cigarettes/day. It appeared

TABLE 6.13

Amount smoked : mean and SD

(cigarettes/day)

	MEN			WOMEN		
	mean	SD	% smokers	mean	SD	% smokers
ENG	4.26	8.17	29	2.18	5.30	20
NEW	9.09	11.45	56	7.28	9.86	48
OLD	9.57	12.07	53	4.62	8.13	33

TABLE 6.14

Cross tabulation for smoking : engaged couples

	HER SMOKING					
	0	0.1-4	5-8	9-12	13-16	17-20
HIS SMOKING cigs/day						
0	45	3	1		1	
0.1-4	3		1			1
5-8						
9-12			1			
13-16	2					1
17-20	6	1			1	
more					2	1

TABLE 6.15

Cross tabulation for smoking : new couples

HER SMOKING

HIS SMOKING cigs/day	0	0.1-4	5-8	9-12	13-16	17-20	21-24	25-28	more
0	34	3	3	3	2	5			
0.1-4	6				1	3			
5-8	4				3		1		
9-12	2			3		1			
13-16	3		1		3	3			
17-20	6		1	1		6			1
21-24									
25-28	1	1				1			2
more	3		1		1	1			1

TABLE 6.16

Cross tabulations for smoking : old couples

HER SMOKING

HIS SMOKING cigs/day	0	0.1-4	5-8	9-12	13-16	17-20	21-24	25-28	29-31	more
0	86	3		2	6	6				1
0.1-4	16	1				1				
5-8	5	1	1							
9-12	7		1	2		2				
13-16	2	3	4		2	1		1		
17-20	15	1	2	6	1	9	1	1	1	
21-24	3									
25-28	5				1					
29-32	5	2		2	1	2			1	
more	5					3				

that the correlation (calculated by the short-cut method of Formula 2.3 on page 39) was remarkably resistant to the choice of level, so we may conclude that 0.1 was an acceptable choice but almost any other would have done.

Tables 6.14 - 6.16 are cross tabulations of the amount smoked in each cohort. The table entries are numbers, not percents. Zero elements have been left blank for clarity. Twenty per day emerges as a watershed: presumably smokers were reluctant to broach the second packet. If we define smoking 20/day as 'heavy' and consider this as a trait in its own right, then several nonsmoking women married heavy smoking men, but few nonsmoking men married a heavy smoking woman. Table 6.17 showed that there was assortative marriage when heavy smoking was treated as an all or none trait, and Table 6.1 showed that there was a correlation even when only smoker x smoker marriages were concerned. So the amount smoked was important as well as the fact of being a smoker.

The amount smoked has been converted to cigarettes, but we ought to look in more detail at methods of smoking. Tables 6.18 - 6.20 were constructed to show the methods used: cigarettes, cigars, and pipes. Some men smoked in more than one way but one method would always predominate. All the women smokers in this survey smoked cigarettes except one woman who occasionally smoked a pipe. Therefore they were tabulated by the amount smoked. Two conclusions may be drawn. Firstly, about 15 - 20% of male smokers predominantly used cigars or pipes. Secondly, of these most married nonsmoking women, and were themselves light smokers.

We would expect smoking to be linked to social class and education, perhaps also to age or other traits. Table 6.21 shows the correlation between smoking and a variety of traits. Only education and social class were consistently linked. It appears that for women

TABLE 6.17

The effect of various cut-off levels on concordance

where: NS.NS = neither smoke; NS.S = woman only smokes; S.NS = man only smokes; S.S = both smoke

CUT OFF LEVEL (cigs/day)

	0.1	0.3	0.9	1.5	4.9	19.5
ENG						
NS.NS	43(63%)	45(66%)	45(66%)	46(68%)	49(72%)	55(81%)
NS.S	5(7%)	6(9%)	6(9%)	4(6%)	4(6%)	2(3%)
S.NS	11(16%)	9(13%)	9(13%)	9(13%)	9(13%)	11(16%)
S.S	9(13%)	8(12%)	8(12%)	7(10%)	6(9%)	0(0%)
r	.390	.378	.378	.347	.380	-.076
SD	.046****	.114***	.114***	.115**	.114***	.123
NEW						
NS.NS	34(30%)	35(31%)	35(31%)	39(35%)	43(38%)	72(61%)
NS.S	16(14%)	16(14%)	17(15%)	15(13%)	17(16%)	12(11%)
S.NS	25(22%)	25(22%)	25(22%)	25(22%)	22(19%)	18(16%)
S.S	38(34%)	37(33%)	36(32%)	34(30%)	31(27%)	11(10%)
r	.282	.282	.263	.301	.304	.257
SD	.091**	.091**	.092**	.091***	.090***	.092**
OLD						
NS.NS	86(39%)	88(40%)	88(40%)	96(43%)	106(48%)	145(65%)
NS.S	18(8%)	19(9%)	19(9%)	16(7%)	16(7%)	12(5%)
S.NS	63(28%)	62(28%)	62(28%)	59(27%)	54(24%)	48(22%)
S.S	55(25%)	53(24%)	53(24%)	51(23%)	46(21%)	17(8%)
r	.311	.302	.302	.349	.365	.250
SD	.064****	.064****	.064****	.063****	.063****	.065***

TABLE 6.18

ENG : Type of Smoking

WOMEN : cigs/day

		NIL	1 - 9	10 - 19	20+	TOT
	NIL	42	3	2		47
	CIGARETTES	10	1	4	3	18
MEN	CIGARS	1		1		2
	PIPE	1				1
	TOT	54	4	7	3	68

TABLE 6.19

NEW : Type of Smoking

WOMEN : cigs/day

		NIL	1 - 9	10 - 19	20+	TOT
	NIL	32	3	7	4	44
	CIGARETTES	18	7	12	20	57
MEN	CIGARS	5		1	1	7
	PIPE	2			1	3
	TOT	55	10	20	26	113

TABLE 6.20

OLD : Type of Smoking

WOMEN: cigs./day

		NIL	1 - 9	10 - 19	20+	TOT
	NIL	90	3	6	6	105
	CIGARETTES	42	14	17	22	95
MEN	CIGARS	11	1			12
	PIPE	9			1	10
	TOT	152	18	23	29	222

it was education rather than social class which determined smoking habits, whilst for men both factors were important. It is also possible that the correlation for smoking varies in different social classes. Table 6.22 shows correlations for smoking when the sample was partitioned in two ways:

i) by husband's social class: IIIN and above versus IIIM and below.

ii) by husband's education: with Highers and above versus O-levels or less.

Ideally we should have partitioned into individual social classes and educational categories, but the numbers were too small for that. The standard deviations were greatly increased and so the significance levels declined. However the values of r in each subgroup remained broadly similar, except for the near zero correlation in the lower social class, low education subgroups of the engaged. This involved only a handful of couples so it seemed reasonable to ignore them. On the basis of these figures there is no reason to suppose that association for smoking was greatly different in individual social classes. (Partition by wife's social and educational categories - not tabulated - produced similar results). Table 6.23 charts smoking by cohort, sex, social class and education. Finally, Table 6.24 shows that the correlation for smoking remained high when corrected for social class and education, and vice versa.

IV. DISCUSSION

The social factors of education, occupation, smoking and religious practice were highly correlated in all three groups of couples ($r > .3$). This implies that each of these factors plays a part in selection of a partner, but they do not tend to converge or diverge thereafter. Partial corrections indicated that these traits were independent of each other, except for education and occupation. This confirms the finding of Warren (1966) that little association for social class remains after education is accounted

TABLE 6.21

Intercorrelation of smoking with other traits

	ENG (n=68)				NEW (n=113)				OLD (n=222)			
	M		F		M		F		M		F	
	r	SD	r	SD	r	SD	r	SD	r	SD	r	SD
Age now	-.038	.123	-.179	.121	-.051	.095	.025	.095	-.020	.067	.095	.067
education	-.271	.118*	-.459	.109****	-.375	.088****	.267	.091**	-.297	.064****	-.190	.066**
neuroticism	-.110	.122	.099	.122	-.025	.095	-.065	.095	.154	.067*	.079	.067
extroversion	.224	.120	-.151	.122	.008	.095	.195	.093*	.127	.067	.039	.068
job when married	.390	.112***	.255	.120*	.367	.091****	.196	.107	.289	.065****	.155	.096
best job	.292	.118 *	.219	.120	.350	.089****	.229	.092 *	.314	.064****	.062	.067
father's job	.028	.124	.034	.123	.199	.093*	.119	.094	.131	.067	.104	.068
religion	-.195	.121	-.220	.120	-.184	.094	-.171	.094	-.169	.066*	-.167	.066*
best height	-.059	.123	-.109	.122	-.149	.094	-.067	.095	-.071	.067	-.037	.067
best weight	.048	.123	-.175	.121	-.201	.093 *	-.029	.095	-.079	.067	-.091	.067

TABLE 6.22

Correlation for smoking in different subgroups in the sample

	(A) by class						(B) by education					
	I + II + IIIN			IIIM + IV + V			with 'Highers' or degree			with 'O' levels or nothing		
	n	r	SD	n	r	SD	n	r	SD	n	r	SD
ENG	58	.545	.112****		NC		54	.651	.105****	14	-.028	.289
NEW	67	.348	.116**	46	.183	.148	53	.159	.138	60	.286	.126*
OLD	132	.288	.084***	90	.274	.103**	78	.273	.110*	144	.296	.080***

NC - not computable, only 10 low social class engaged couples

TABLE 6.23

Amount smoked by cohort, class, sex and education.

Group	TOTAL	NIL	½-9	10-19	20-29	30-39	40+
a) ENGAGED							
All men	68	51	2	4	11		
All women	68	54	7	5	2		
Men, Class I-IIIN	58	46	2	1	9		
Men, Class IIIM-V	10	5	0	3	2		
Women, Class I-IIIN	65	52	7	5	1		
Women, Class IIIM-V	3	2	0	0	1		
Men, Educ ⩾ Highers	54	42	2	2	8		
Men, Educ ⩽ O level	14	9	0	2	3		
Women, Educ ⩾ Highers	56	48	4	4	0		
Women, Educ ⩽ O level	12	6	3	1	2		
b) NEW							
All men	113	51	17	16	20	5	4
All women	113	60	11	19	19	2	2
Men, Class I-IIIN	67	38	9	8	9	2	1
Men, Class IIIM-V	46	13	8	8	11	3	3
Women, Class I-IIIN	95	54	8	15	14	2	2
Women, Class IIIM-V	18	6	3	4	5	0	0
Men, Educ ⩾ Highers	53	34	6	6	5	1	1
Men, Educ ⩽ O level	60	17	11	10	15	4	3
Women, Educ ⩾ Highers	58	38	7	4	6	1	2
Women, Educ ⩽ O level	55	22	4	15	13	1	0
c) OLD							
All men	222	107	24	26	43	15	7
All women	222	150	20	23	26	2	1
Men, Class I-IIIN	131	77	17	15	19	2	1
Men, Class IIIM-V	91	30	7	11	24	13	6
Women, Class I-IIIN	174	122	13	19	19	0	1
Women, Class IIIM-V	48	28	7	4	7	2	0
Men, Educ ⩾ Highers	78	48	9	10	10	1	0
Men, Educ ⩽ O level	144	59	15	16	33	14	7
Women, Educ ⩾ Highers	79	58	10	6	5	0	0
Women, Educ ⩽ O Level	144	92	10	17	21	2	1

TABLE 6.24

Partial Correction of Trait A for Trait B

Trait A	Trait B	ENG (n=68) r	ENG (n=68) SD	NEW (n=113) r	NEW (n=113) SD	OLD (n=222= r	OLD (n=222= SD
church	bestjob	.517	.107****	.415	.088****	.539	.057****
bestjob	church	.350	.117*	.552	.081****	.473	.060****
church	school	.507	.108****	.400	.089****	.544	.057****
school	church	.467	.111****	.583	.079****	.636	.052****
church	smoking	.493	.109****	.394	.089****	.549	.057****
smoking	church	.409	.114***	.287	.093***	.306	.064****
bestjob	school	.219	.122	.193	.094*	.058	.068
school	bestjob	.347	.117**	.255	.094*	.274	.065****
bestjob	smoking	.219	.122*	.485	.084****	.472	.060****
smoking	bestjob	.361	.117**	.229	.093*	.283	.065****
school	smoking	.418	.114***	.531	.081****	.619	.053****
smoking	school	.368	.116**	.222	.093*	.287	.065****
bestjob	patjob	.241	.122*	.417	.087****	.345	.064****
patjob	bestjob	.150	.125	.144	.095	.115	.068
school	patjob	.429	.114***	.468	.085****	.511	.059****
patjob	school	.248	.122*	.156	.095	.110	.068

Table 6.25

Further corrections of social traits for religion practice

Subset	ENG			NEW			OLD		
	n	r	SD	n	r	SD	n	r	SD
<u>Religious men</u>									
bestjob	29	.293	.184	19	.496	.211*	64	.371	.118**
school	29	.459	.171****	19	.391	.223	64	.541	.107****
patjob	29	.245	.187	19	.164	.239	63	.212	.125
smoking	29	.572	.158***	19	.119	.241	64	.282	.122*
<u>Nonreligious men</u>									
bestjob	39	.323	.156*	92	.538	.089****	158	.510	.069****
school	39	.473	.145**	92	.620	.083****	158	.666	.060****
patjob	38	.372	.155*	91	.395	.097****	155	.307	.077****
smoking	39	.375	.152*	92	.327	.100**	158	.336	.075****
<u>Religious women</u>									
bestjob	36	.125	.170	37	.590	.136****	102	.380	.093****
school	36	.435	.154**	37	.500	.146***	102	.550	.084****
patjob	36	.279	.165	37	.358	.158*	100	.265	.097**
smoking	36	.530	.145***	37	.310	.161	102	.362	.093***
<u>Nonreligious women</u>									
bestjob	32	.450	.163**	74	.488	.103****	120	.549	.077****
school	32	.522	.156***	74	.616	.093****	120	.703	.065****
patjob	31	.379	.172*	73	.299	.113**	118	.325	.088***
smoking	32	.385	.169*	74	.346	.111**	120	.252	.089**

for. However, we have been using parametric methods throughout, and at this stage it is worth rehearsing the arguments for doing so. Parametric analysis has greatly facilitated our examination of social traits but it represents an innovation which may affright the statistically orthodox. It is hoped that the following exposition will assure investigators in the social sciences that the domain of parametric methods may be set wider than they had hitherto considered, and will encourage them to further statistical experiments with their own data.

Firstly, social class itself is not very far from Normality, by the lax standards of real-life biological data. Society at large is generally held to be "onion-shaped" with the bulk in the middle and a tail at each end. Both class and education may be thought of as representing underlying abilities such as intelligence which there is every reason to think of as "Normal". However this argument clearly does not hold for religion or smoking habits, which are anything but Normal.

Secondly, the calculation of r is highly robust with respect to shape of distribution, even in the extreme "all or none" classification of, say, religion.

Thirdly, the number of observations, n, is very large. With nonparametric methods n is generally considered to be adequate in the range 4 to 25, and large from 26 to 100; anything larger is an indication for a parametric method. With the numbers in the present study the value of r approximated closely to that of Spearman's (nonparametric) correlation coefficient, r_s. This is demonstrated in Table 6.26, where social class, education, father's social class and smoking show good agreement between the two methods. For religion they are exactly equal; it has been left as an exercise for the reader to prove that

TABLE 6.26

Comparison of parametric (r) and nonparametric (r_ρ) inter-spouse correlation coefficients for social factors

	ENG		NEW		OLD	
	r	r_ρ	r	r_ρ	r	r_ρ
BESTJOB	.298*	.279*	.553****	.526****	.488****	.479****
SCHOOL	.478****	.447****	.596****	.593****	.640****	.611****
PATJOB	.310**	.369**	.327***	.364****	.299****	.312****
SMOKING	.431***	.436****	.311***	.343***	.314****	.338****
CHURCH	.515****	.515****	.415****	.415****	.551****	.551****

in the case of all-or-none coding, Formula 2.4 for r_s is mathematically identical to the short-cut Formula 2.3 for r.

Fourthly, the significance levels of the results are clear cut. The interspouse correlations for social factors are evident simply by inspecting the data. The Null Hypothesis, that this so-called assortative mating is merely the result of chance, is not credible. The hypothesis we really wish to test is whether true AM for one trait can explain away the correlation for another, and it is here that a more serious objection to the methodology can be raised. The calculation of partial correlations is very sensitive to non-Normality and unfortunately there is no equivalent nonparametric test.

Confirmation of results must therefore be sought by partitioning the sample, eg to correct correlations of social class for education one must calculate r_s for social class in each educational subgroup. Note that this partitioning must be fine and painstaking. It was found by trial and error that it was not adequate to simply partition into the five levels of the man's education, as the woman's education still influenced the results. It was necessary to consider each of the 25 combinations, and frame an appropriate series of commands, thus:

```
SELECT ONLY HIS SCHOOL = 1  AND HER SCHOOL = 1
NONPARM HIS BESTJOB V HER BESTJOB 1 0.5 5  1   0.5  5
SELECT ONLY HIS SCHOOL = 1  AND HER SCHOOL = 2
```

and so on. In order to boost the numbers in each cell it was decided to combine all three cohorts. Nevertheless some cells remained empty and were ignored, whilst others were inquorate and were combined with adjacent cells.

In all this process generated 13 subgroups of social class, 12 of education, 16 of father's social class, 9 of smoking, and 4 of religion. In each subgroup r_s was calculated for all

TABLE 6.27

Correction of Trait A for Trait B - comparison of parametric and nonparametric methods.
All three cohorts combined. (Parametric results in brackets)

TRAIT A	UNCORRECTED	TRAIT B: BESTJOB	SCHOOL	PATJOB	SMOKING	CHURCH
BESTJOB	.502 (.509)	-	.163 (.118)	.311 (.346)	.447 (.463)	.471 (.501)
SCHOOL	.648 (.650)	.318 (.322)	-	.472 (.502)	.610 (.609)	.629 (.642)
PATJOB	.362 (.362)	.157 (.163)	.155 (.162)	-	.340 (.339)	.348 (.354)
SMOKING	.364 (.331)	.324 (.276)	.310 (.273)	.342 (.311)	-	.364 (.313)
CHURCH	.517 (.517)	.507 (.508)	.510 (.508)	.492 (.514)	.504 (.502)	-

the social traits other than the one defining the subgroup. This r_ς was then weighted by n and summed to give a rank correlation coefficient corrected for the influence of each trait. These results are presented in Table 6.27. For ease of comparison the parametric equivalent is given below each result in brackets. Also given are the uncorrected results for the combined cohorts. (These are slightly higher than those obtained by summing the engaged, newly wed and old results previously tabulated. In effect we have ignored the trait "their cohort" to concentrate on other factors.)

It will be seen that the agreement between the two methods is close. To save space standard deviations are not tabulated; in fact all lay between 0.05 and 0.04. That means that all correlations were significant at the level $p \geqslant 0.0001$ with three important exceptions. Father's job, corrected for either bestjob or education, only reached the $p \geqslant 0.01$ level with $r_\varsigma = .155 \pm 0.049^{**}$. It looks as though that correlation, which was .362 uncorrected, would be reduced considerably further if we could have corrected for both bestjob and education simultaneously. Social class, corrected for education, was still highly significant by both parametric and non-parametric methods ($r_\varsigma = .163 \pm .049^{***}$) although greatly reduced. It was also reduced by correction for father's job, but that might have been a knock-on effect of education. Table 6.27 demonstrates that religion and smoking were independent both of each other and of the other social traits. But bestjob, school, father's job and cohort remained to some extent intertwined, and it seemed desirable to try to unravel the strands a little further.

One approach might be to attempt partitioning by two or more traits simultaneously. It could be done manually thus:

```
SELECT ONLY HIS SCHOOL = 1 AND HER SCHOOL = 1 AND THEIR COHORT = 'E'
SELECT ABOVE AND HIS PATJOB = 1 AND HER PATJOB = 1
NONPARM HIS BESTJOB V HER BESTJOB 1 0.5 5 1 0.5 5
```

Alternatively one could automate the subgrouping, modifying the program to create a table in a hyperspace whose dimensions were the controlling traits, and to merge cells as required. One would then have a truly multivariate nonparametric function, accessible through a new command:

```
MULTIVARS BESTJOB V PATJOB AND SCHOOL AND COHORT
```

However instead of embarking on these complexities we might learn something from simpler analysis. The bulk of the association for bestjob was found to be coming from the lowest educational category where his school < 1.5 and her school $<$ 1.5: n = 102 $r_s = .225 \pm .097$*. It seemed that this category, which could not be further subdivided by education, was obscuring genuine differences in talent. It was mentioned on page 129 that many older people had never had the opportunity of higher education, and presumably they would still assort by social class. If that were true, the correlation for bestjob might vanish in a younger cohort.

To test this the corrections for bestjob, school and father's job were repeated with only those whose agenow < 35 included - i.e. those born after 1943. The correlation for education was not affected, but the association for bestjob and father's job both dwindled further (Bestjob: $r_s = .139 \pm .057$* and $.279 \pm .036$**** corrected for education and father's job respectively; patjob: $r_s = .162 \pm .057$** and $.143 \pm .057$* corrected for bestjob and education respectively). We may therefore conclude that in a post-war sample, education is almost (but not quite) sufficient to account for matching in bestjob and father's job, but that this is not true of older couples.

There is also the question of differences between cohorts, which were glossed over above. Most of these are trivial, but again social class and education stand out. During the earliest runs of the program DANTE, before many revisits had been made, it was noticed that bestjob was not correlated in old couples when corrected for education but was significant at the 1% level in new and engaged couples. As the months went by and revisits continued, this significance dwindled away almost completely. Something was happening in the first few months of marriage. The simplest explanation is that many young people were having difficulty finding work to match their education. There were several recent graduates who at the first visit were van drivers, baggage porters, and the like. By the following year they were lab technicians, trainee managers, etc. (Thus, they were seeking Class II jobs: the professionally qualified had fewer problems.) It appears that bestjob was slowly moving to mirror education but in old couples the process was complete.

Of those potential members of Class II, probably the women were having the most difficulty as their career opportunities were comparitively restricted. For instance, in Table 6.2 we see that only one man in Class I chose a woman in a class as low as IIIN. Inspection of this couple's original file (E54) showed that she was actually better educated than he was, having a degree. But they had moved to the far North, where the available careers for a woman arts graduate would be teaching or nothing.

Given long enough, one would expect male occupation to match very closely to education as men rose to their innate level of competence. But with women this process would be curtailed by deciding to start a family. There would also be some connection with age as seniority brought promotion. However promotion ladders within a given career structure seldom cross social class boundaries. A student nurse will remain in Class II even if she becomes a matron or penetrates the higher echelons beyond, and entrance to the professions is guarded by university intake. Only the less common changes, from shop-floor or filing office to manager's desk, are reflected in elevation of class.

Coleman (1977) had found assortative marriage for social class to be much stronger in Class I than III, in surveys carried out both by himself (Coleman 1973) and by Glass (1954). However his measure was not 'r' but the 'index of association' - the ratio of observed to expected marriages in a given category. Suppose that 5% of both men and women are in Class I, then 5% * 5% = 0.25% of marriages would be expected to be assortative class I x class I. If 2% are observed, the index is 8. But the smaller the category the larger the index can be. If Class I practised strict assortative marriage, the index would be 5% / 0.25% = 20. If Class III also practised strict assortative marriage and comprised 50% of the population, the index would only be 50% / (50% * 50%) = 2. Therefore this 'difference' in behaviour is largely due to the onion-shaped distribution of social classes. One would need to supplement the index with some further statistic to test the significance of a given deviation from unity.

Smoking is among those traits which seem to be important in selecting a partner. This is hardly surprising, as the offer of a cigarette is a standard opening gambit in the courtship ritual. Smoking habits were interlinked with the other psycho-social factors mentioned but Tables 6.25 and 6.24 suggest that they were assorting independently with r in the range 0.25 - 0.35. Of course there is always the possibility that this was merely reflecting some other factor not examined such as IQ. But if IQ was the underlying cause this should have been revealed by correction for education.

Other mechanisms strengthening this correlation are, firstly that nonsmokers may participate more in active sport, where they are less likely to socialise with smokers. Since we have no information on exercise patterns, and have found no correlation for fitness as reflected by pulse and blood pressure, we can not test this hypothesis. Secondly, even quite light smokers acquire a distinctive smoky odour on their clothes and breath. Other things being equal, a nonsmoker would probably prefer to kiss a nonsmoker than a smoker. Thirdly, a nonsmoker who is sanctimonious or moralistic about smoking will prove intensely irritating to a smoker.

The smoking habits of the parents of these couples were not recorded and their own children were very young. Therefore it can not be shown whether assortative marriage was influencing the frequency of the trait. Since smoking is learnt partly by parental example, this influence must be increased if marriages are concordant, as they tend to be.

It had been expected that at this stage a correction would be made for the social class and other biases in response rate detailed in Chapter 2. However, although many of the traits examined were associated with social class, there was no evidence of different behaviour in different social classes. For instance, there are far more smokers in the lower social classes but smokers prefer each other in all classes, and no correction appears necessary.

Also intended was a correction for those couples ascertained as "mistakes", and here again no such correction was necessary. The 23 "mistakes" had been redistributed as 7 new and 16 old couples so their effect on either cohort was small. Re-runs on the new and old cohorts with "mistakes" excluded changed the third place of decimals in the value of r but rarely its significance level. (These re-runs are not tabulated.) It seems safe, therefore, to ignore the effects of these couples on the final results.

CHAPTER 7.

Conclusions

CHAPTER SEVEN

The correlations between spouses for the principal traits investigated in this study are summarised in Table 7.1. The most important factors emerge as age at marriage, height, neuroticism, education, religious practice and smoking habits. These factors all appear to assert independently, whereas the correlation for social class turned out to be explicable in terms of education. A search was made for evidence of convergence or divergence of traits with time but there was little evidence of this, either on comparing the same couples a year apart or by comparing engaged couples with newly weds and those married five years or more. Because there were few separations in the sample it was not possible to demonstrate whether lack of correlation led to marriage break-up.

What might be the biological role of assortative marriage? It is such a ubiquitous phenomenon that it must surely be of some benefit at the species as well as individual level. In the foregoing chapters reasons were suggested why homogamy might be to the benefit of the individuals concerned. If this does in fact lead to harmonious partnership, and to higher fertility, then no further explanation need be sought. Yet, for example, height was more highly correlated than intro/extroversion, though one might have predicted that the latter was more conducive to harmony.

Even experiments specifically designed to demonstrate biological consequences of assortative marriage have failed to show them. For instance, the Canadian lesser snow goose practises homogamy for a shade of down. Because the geese are usually monogamous, and return to their birthplace after migration, it is possible to test their biological fitness in ringing experiments. No difference in fitness could be found between geese that chose like partners and those that chose unlike (Cooke 1979). Actually, perhaps the most important aspect of assortative mating is overlooked by this experiment - homogamy for flock and migration pattern. Perhaps the explanation is to be found in the effects on species variance and response to selection demonstrated by McBride et al (1963). When there is assortative marriage the offspring of any given partnership show reduced variance, but the offspring of the entire sample show increased variance. This preserves variety and flexibility within the species, which is therefore better able to respond to selection pressure. However, in the only genetic factor, height, it could not be shown that height influenced fertility.

Psychosocial traits which are not genetic but nevertheless familial, have generally shown stronger correlations than physical traits. Although it is impossible to prove by controlled experiments, by analogy with inherited biological traits the population benefits of assortative marriage may be the same for psychosocial traits. Societies expand, evolve, mutate, subdivide and senesce as do species. Flexible societies prosper, rigid ones decline. Assortative marriage, by promoting diversity and hence flexibility, may lie at the roots of social evolution.

TABLE 7.1

Summary of principal correlations

	ENG (n=68)		NEW (n=113)		OLD (n=222)	
	r	SD	r	SD	r	SD
Age	.456	.110****	.903	.041****	.889	.031****
Bestht	.301	.117*	.170	.094	.304	.064****
Neuro	.247	.119*	.330	.090***	.237	.066***
Extro	.097	.123	.290	.091**	.039	.068
Bestjob +	.219	.122	.193	.094*	.058	.068
Education +	.347	.117**	.255	.094**	.274	.065****
Religion	.515	.106****	.415	.087****	.551	.056****
Smoking	.431	.111****	.311	.090***	.314	.064****

+ Best job and Education corrected for each other

APPENDIX

Appendix A: Eysenck's Personality Inventory

Here are some questions regarding the way you behave, feel and act. After each question there is a 'YES' and '?' and 'NO'. Try and decide whether 'YES' or 'NO' represents your usual way of acting and feeling, then put a circle round the 'YES' or 'NO'. If you find it absolutely impossible to decide, put a circle round the '?', but try not to use the '?' too often.

1.	Are you happiest when you get involved in some project that calls for rapid action?	YES	?	NO
2.	Do you sometimes feel happy, sometimes depressed, without any apparent reason?	YES	?	NO
3.	Does your mind often wander while you are trying to concentrate?	YES	?	NO
4.	Do you usually take the initiative in making new friends?	YES	?	NO
5.	Are you inclined to be quick and sure in your actions?	YES	?	NO
6.	Are you frequently lost in thought even when supposed to be taking part in a conversation?	YES	?	NO
7.	Are you sometimes bubbling over with energy and sometimes very sluggish?	YES	?	NO
8.	Would you rate yourself as a lively individual?	YES	?	NO
9.	Would you be very unhappy if you were prevented from making numerous social contacts?	YES	?	NO
10.	Are you inclined to be moody?	YES	?	NO
11.	Do you have frequent ups and downs in mood, with or without apparent cause?	YES	?	NO
12.	Do you prefer action to planning for action?	YES	?	NO

PROFESSOR ALAN E.H. EMERY
M.D., Ph.D., D.Sc., F.R.C.P. (E.), M.F.C.M.,
F.R.S. (E.).

University of Edinburgh

DEPARTMENT OF HUMAN GENETICS
WESTERN GENERAL HOSPITAL
EDINBURGH EH4 2XU
TELEPHONE: 031–332 2471

APPENDIX B 1

This letter sent to
engaged and new couples

Dear

We are undertaking some medical research in which we wish to measure similarities between couples for certain characteristics such as height and weight. These couples include those who are married, and those who have recently become engaged. To get meaningful results we need to study as many couples as possible.

Would it be at all convenient for me to visit you both, to take the necessary measurements? In any event I would be most grateful if you could return the enclosed card, duly completed.

Thank you very much for your cooperation which is greatly appreciated.

Yours sincerely,

Dr. G.C. Sutton, M.Sc., M.B., Ch.B.,
Medical Research Fellow

PROFESSOR ALAN E.H. EMERY
M.D., Ph.D., D.Sc., F.R.C.P.(E.), F.R.S.(E.)

Universitas Edinburgh

DEPARTMENT OF HUMAN GENETICS
WESTERN GENERAL HOSPITAL
EDINBURGH. EH4 2HU
TELEPHONE: 031-332 2471

APPENDIX B 2

This letter sent to old couples

Dear

We are undertaking some medical research in which we wish to measure similarities between people for certain characteristics such as height and weight. To get meaningful results we need to study as many people as possible.

Would it be at all convenient for me to visit you to take the necessary measurements? In any event I would be most grateful if you could return the enclosed card, duly completed..

Thank you very much for your cooperation, which is greatly appreciated.

Yours sincerely,

Dr. G.C. Sutton, M.Sc., M.B., Ch.B.
Medical Research Fellow.

APPENDIX B 3

Dr. Graham Sutton,
Dept. of Human Genetics
Western General Hospital
Edinburgh,
EH4 2XU

Please delete as appropriate:-

(a) It will be convenient for you to call

on .... at ....

(if a different time suits you better, please write that in).

or (b) It will not be convenient for you to call.

Please correct the address if need be, and give a phone number if possible.

PROFESSOR ALAN E.H. EMERY
M.D., Ph.D., D.Sc., F.R.C.P. (E.), M.F.C.M., F.R.S. (E.).

University of Edinburgh

DEPARTMENT OF HUMAN GENETICS
WESTERN GENERAL HOSPITAL
EDINBURGH EH4 2XU
TELEPHONE: 031-332 2471

APPENDIX B 4

Dear

Some weeks ago you kindly allowed me to visit you in connection with some medical research. You may recall that at that time I was missing a piece of equipment to do one further measurement, namely "skin fold thickness". This equipment has now, at last, arrived, and I wonder if it would be possible for me to call to take the readings. This is a simple (and painless) procedure which should take me less than 5 minutes to complete.

Could you, as before, complete and return the enclosed postcard, so that I will know when and if I can call?

Many thanks.

Yours sincerely,

Dr. G.C. Sutton

PROFESSOR ALAN E.H. EMERY
M.D., Ph.D., D.Sc., F.R.C.P. (E.), M.F.C.M., F.R.S. (E.).

University of Edinburgh

DEPARTMENT OF HUMAN GENETICS
WESTERN GENERAL HOSPITAL
EDINBURGH EH4 2XU
TELEPHONE: 031-332 2471

APPENDIX B 5

Dear

It is now roughly 12 months since I visited you in connection with my project to measure similarity between couples. I now wish to visit and measure each couple one last time, to discover whether they are growing more or less similar as time goes by. The measurements involved are much the same as last year, although the visit should take less time as I do not intend to repeat all of them.

I would be grateful if you could, as before, complete and return the postcard, so that I will know when and if I should call.

Many thanks.

Yours sincerely,

Dr. G.C. Sutton

BIBLIOGRAPHY

ADAMS M.S. (1969)
Genetic Consequences of Cultural Adaptation
Medical Clinics of North America 53 (4) 977-990

BECKMANN L. AND ELSTON R. (1962)
Assortative Mating and Fertility.
Acta Genetica et Statistica Medica 12 117-122

BLOCK N. AND DWORKIN G. (1977)
The IQ Controversy.
Quartet Books, London.

BOWERMAN C.E. (1953)
Assortative Mating by Previous Marital Status Seattle 1939 - 1946
American Sociology Review 18 171-177

BURGESS E.W. AND WALLIN P. (1944)
Homogamy in Personality Characteristics.
Journal of Abnormal and Social Psychology 39 475-481

BURT C. AND HOWARD M. (1956)
The Multifactorial Theory of Inheritance and Its Application to Intelligence.
British Journal Of Statistical Psychology 9 95-131

CAVALLI-SFORZA L. AND BODMER W.F. (1971)
The Genetics of Human Populations,
W.H. Freeman & Co., San Francisco.

COLEMAN D.A. (1977)
Assortative Marriage in Britain.
Chap. 2 of 'Equalities and Inequalities in Family Life'.
Editors: Chester R. and Peel J. Academic Press, London.

COOKE F. (1979)
How Do We Choose Our Mate?
New Scientist 81 325-327

CROW J.F. AND FELSENSTEIN J. (1968)
The Effect of Assortative Mating on the Genetic Composition of a Population.
Eugenics Quarterly 15 (2) 71-84

DOMINIAN J. (1979) Marital Pathology (11 Papers, British Medical Journal 2).
A. Health & Marital Breakdown 424-425
B. Definition & Extent of Marital Pathology 478-479
C. Social Factors and Marital Pathology 531-532
D. Choice of Partner 594-596
E. First Phase of Marriage 654-656
F. Second Phase of Marriage 720-722
G. Third Phase of Marriage 781-783
H. Marriage & Psychiatric Illness 854-855
I. Management: Basic Counselling 915-916
J. Management: Psychodynamics 987-989
K. Management: Sexual Counselling 1053-1054

DREWERY J. AND RAE J.B. (1969)
A Group Comparison of Alcoholic and Non-Alcoholic Marriages Using the Interpersonal Perception Technique.
British Journal of Psychiatry 115 287-300

DURNIN J.V.G.A. AND WOMERSLY J. (1974)
Body Fat Assessed From Total Body Density and its Estimation from Skinfold Thickness: Measurements on 481 Men and Women Aged from 16 to 72 Years
British Journal of Nutrition 32 77-97

ECKLAND B.K. (1968)
Theories of Mate Selection.
Eugenics Quarterly 15 (2) 71-84

EDINBURGH REGIONAL COMPUTING CENTRE (1970)
Edinburgh IMP Language Manual.
Edinburgh University.

EMERY A.E.H. (1976)
Methodology in Medical Genetics.
Churchill Livingstone, Edinburgh.

EYSENCK H.J. (1960)
The Structure of Human Personality
2nd Edn. P44-50
Methuen, London

EYSENCK H.J. (1962)
Response Set, Authoritarianism and Personality Questionaires.
British Journal of Social and Clinical Psychology 1 20-24

EYSENCK H.J. AND EYSENCK S.B.G. (1964)
Manual of the EPI.
University of London

FRASER G.R. (1976)
The Causes of Profound Deafness in Childhood.
Johns Hopkins, Baltimore & London.

GARROW J.S. (1979)
Weight Penalties .
British Medical Journal 2 1171-1172

GLASS D.V. (1954) (EDITOR)
Social Mobility in Britain.
Routledge and Kegan Paul, London.

GRIFFITHS R.W. AND KUNZ P.R. (1973)
Assortative Mating - A Study of Physiognomic Homogamy.
Social Biology 20 448-453

GUTTMAN R. (1970)
Parent-Offspring Correlations in the Judgement of Visual Number.
Human Heredity 20 57-65

HARRISON G.A., GIBSON G.B., AND HIORNS R.W. (1976)
Assortative Mating for Psychometric, Personality and Anthropometric Variations in a Group of Oxfordshire Villages.
Journal of Biosocial Science 8 145-153

JOHNSON B.C., EPSTEIN F.H., AND KJELSBERG M.O. (1965)
Distributions and Familial Studies of Blood Pressure and Serum Cholesterol Levels in a Total Community - Tecumseh, Michigan.
Journal of Chronic Diseases 18 147-160

JOHNSON R.C., DEFRIES J.C., WILSON J.R., McCLEARN G.E., VANDENBERG S.G., ASHTON G.C., MI M.P., AND RASHAD M.N. (1976)
Assortative Marriage for Specific Cognitive Abilities in Two Ethnic Groups ,
Human Biology 48 343-352

JOHNSTON F.E. (1970)
Phenotypic Assortative Mating Among the Peruvian Cashinahua.
Social Biology 17 37-42

KISER C.V. (1968)
Assortative Mating by Educational Attainment in Relation to Fertility.
Eugenics Quarterly 15 (2) 98-112

KOZLOWSKI L.T., HERMAN C.P., AND FRECKER R.C. (1980)
What Researchers Make of What Cigarette Smokers Say: Filtering Smokers' Hot Air.
Lancet 1 699-700

KREITMAN N. (1964)
The Patient's Spouse.
British Journal of Psychiatry 110 159-173

LUCKEY E.B. (1960)
Marital Satisfaction and Congruent Self Spouse Concepts.
Social Forces 39 153-157

MCBRIDE G. AND ROBERTSON A. (1963)
Selection Using Assortative Mating in Drosophila Melanogaster.
Genetics Research 4 356-369

MASCIE-TAYLOR C.G. AND GIBSON J.B. (1979)
A Biological Survey of a Cambridge Suburb: Assortative Marriage for I.Q. and Personality Traits.
Annals of Human Biology 6 (1) 1-16

MITTON J.B. (1975)
Fertility Differentials in Modern Societies Resulting in Normalizing Selection for Height.
Human Biology 47 189-200

OVENSTONE I.M.K. (1973)
The Development of Neurosis in the Wives of Neurotic Men.
Part I: Symptomatology and Personality.
British Journal of Psychiatry 122 35-43

POMERAT C.M. (1936)
Homogamy and Infertility.
Human Biology 8 19-24

REGISTRAR GENERAL SCOTLAND
Population and Vital Statistics Annual Report Part 2 (1972 VOL 118; 1977 VOL 123).

ROBERTS D.F. (1977)
Assortative Mating in Man: Husband/Wife Correlations in Physical Characteristics.
Bulletin of the Eugenics Society Suppl. 2, Sept. 1977 p 1-45

SIEGEL S. (1956)
Non-parametric Statistics for the Behavioral Sciences.
McGraw-Hill, New York.

SMITH M. (1946)
A Research Note on Homogamy of Marriage Partners in Selected Physical Characteristics.
American Sociology Reviews 11 226-228

SNEDECOR G.W. AND COCHRAN W.G. (1967)
Statistical Methods.
6th Edn. Iowa State, Ames, Iowa.

SPUHLER J.N. (1967) (EDITOR & AUTHOR) Mating Patterns.
P. 241-268 of 'Genetic Diversity and Human Behavior'
Viking Foundation Publications in Anthropology 45, Aldine, Chicago.

ANNALS OF HUMAN BIOLOGY, 1980, VOL. 7, NO. 5, 449–456

Assortative marriage for smoking habits

G. C. SUTTON

University Department of Human Genetics, Western General Hospital, Edinburgh

Received 5 November 1979; revised 14 April 1980

Summary. Assortative marriage for smoking habits was studied in 68 engaged couples, 112 newly wed couples, and 223 couples married for six years or more. All three groups showed a fairly strong correlation between couples for smoking, $r = 0{\cdot}332 \pm 0{\cdot}047$ (mean + SD). Smoking habits were highly intercorrelated with social class and education but the correlation for smoking remained highly significant, even when correction was made for these factors. Because this correlation was present even in those who have not yet married, smoking habits may well be an important factor in choosing a marriage partner.

1. Introduction

'Assortative marriage' means the tendency for people to choose spouses who resemble them more closely than would be expected by chance. Assortative marriage is known to occur for a wide variety of physical, psychological and social traits (Spuhler 1968, Roberts 1977, Vandenberg 1972, Coleman 1977). Surprisingly, smoking habits appear to have escaped attention hitherto. Information on assortative marriage would be of interest in a number of ways. To what extent are smoking habits a factor in selection of a prospective partner? What happens when partners have dissimilar smoking patterns—do they tend to converge, or to level up or down? What influence might this have on the smoking habits of the offspring?

The results are therefore presented from a survey of assortative marriage in Edinburgh, and interpreted in the context of closely related variables such as education and social class.

2. Materials and methods

Seven hundred and eighty-three couples married in 1972 ('old' couples) were traced from the Register of Marriages in Scotland for that year. 538 couples married in October to December 1977 ('new' couples) were ascertained from Registry Office notices and newspaper announcements. Three hundred and fifty engaged couples were traced from newspaper announcements.

All couples were sent a letter briefly explaining the nature of the study and inviting them to take part. A prepaid postcard was enclosed for the reply. Eventually 223 'old' couples, 112 'new' couples, and 68 engaged couples, who had agreed to take part and who were still resident in Edinburgh, were visited at home by myself during 1978 and 1979. Thus the 'new' couples had been married for a few months and the 'old' couples for about six years at the time of visiting. The groups to which couples were finally allocated do not correspond exactly to the mode of ascertainment. For example, a number of engaged couples had already married before they could be visited.

A variety of physical, psychological and social traits were measured on each partner. Those relevant to the present report are date of birth, length of marriage, occupation, education, religious practice, smoking habits, personality factors and father's occupation.

Occupations were categorized into social classes according to the Registrar General's Catalogue 1970. Because some people could not be categorized (housewives, the retired, the unemployed or disabled), the composite parameter 'best job' was created. This means the most recent classifiable job the person has had; if none, 'best job' is defined as the father's occupation.

Age was taken as the age at marriage.

Education was scored by qualifications in most cases. However, those who had left school before O-levels, etc., were introduced were scored by school-leaving age.

For the present analysis, religious practice was simply scored as 1 or 0 according to whether the person actively practised any religion. The definition of 'active practice' was left to the individual. It seemed invidious to lay down fixed criteria such as frequency of church-going.

Smoking habits were scored as number of cigarettes smoked per day. As relatively few people smoked pipes or cigars no separate analysis was made of them, but the amount simply converted to an equivalent number of cigarettes. Thus 1 ounce (28 g) of pipe tobacco per week would be scored as 4/day. Those smoking less than 3/week were considered non-smokers.

Personality factors of extroversion and neuroticism were measured by a short form of Eysensk's Personality Inventory (Eysenck 1960).

The IMP computer program DANTE (Sutton, unpublished) was developed to handle the large volume of data collected.

3. Results

Table 1 shows that smoking habits were highly correlated in all three groups, the mean correlation (weighed by n) being $r = 0{\cdot}332 \pm 0{\cdot}047$. The differences between the groups and sexes are presented in table 2, the main finding being that fewer engaged

Table 1. Overall correlation and concordance for smoking.

			Male-female		Non-smoker (NS) v. smoker (S)	
	N	r SD	NS–NS	NS–S	S–NS	S–S
Engaged	68	$0{\cdot}429 \pm 0{\cdot}111$****	45 (66%)	6 (9%)	9 (13%)	8 (12%)
New	112	$0{\cdot}319 \pm 0{\cdot}090$***	35 (31%)	16 (14%)	25 (23%)	36 (32%)
Old	223	$0{\cdot}309 \pm 0{\cdot}064$****	88 (39%)	19 (9%)	62 (28%)	54 (24%)
Weighted mean: $r = 0{\cdot}332 \pm 0{\cdot}047$****						

In all the tables: * indicates $1\% \leq P < 5\%$

** $0{\cdot}1\% \leq P < 1\%$

*** $0{\cdot}01 \leq P < 0{\cdot}1\%$

**** $P < 0{\cdot}01\%$

Table 2. Amount smoked

	Men			Women		
	Mean	SD	% smokers	Mean	SD	% smokers
Engaged	4·25	8·17	25	2·18	5·30	21
New	9·14	11·49	55	7·16	9·83	46
Old	9·49	12·07	52	4·69	8·18	33

Table 3. Intercorrelation of smoking with other traits.

	Engaged		New		Old	
	M	F	M	F	M	F
	r SD	r SD	r SD	r SD	r SD	r SD
Age	−0·038 ± 0·123	−0·191 ± 0·121	−0·051 ± 0·095	0·016 ± 0·095	−0·068 ± 0·067	0·047 ± 0·067
Education	−0·271 ± 0·118*	−0·459 ± 0·109****	−0·381 ± 0·088****	−0·257 ± 0·092**	−0·300 ± 0·064****	−0·196 ± 0·066**
Neuroticism	−0·111 ± 0·122	0·099 ± 0·122	−0·028 ± 0·095	−0·083 ± 0·095	0·154 ± 0·067*	0·092 ± 0·067
Extroversion	0·224 ± 0·120	−0·151 ± 0·122	0·006 ± 0·095	0·180 ± 0·094	0·129 ± 0·067	0·052 ± 0·067
Social class	0·390 ± 0·113***	0·271 ± 0·120*	0·369 ± 0·092****	0·180 ± 0·108	0·294 ± 0·065****	0·145 ± 0·096
Father's job	0·029 ± 0·124	−0·034 ± 0·123	0·202 ± 0·094*	0·127 ± 0·095	0·128 ± 0·067*	0·092 ± 0·068
Best job	0·323 ± 0·116**	0·233 ± 0·120	0·353 ± 0·089****	0·220 ± 0·093*	0·321 ± 0·064****	0·051 ± 0·067
Religion	−0·195 ± 0·121	−0·216 ± 0·120	−0·182 ± 0·095	−0·164 ± 0·095	−0·165 ± 0·066*	−0·173 ± 0·066**

people smoked. From table 2 we can also see that if people married at random with respect to smoking habits, then both partners would be expected to be smokers in 5% of engaged couples, 25% of 'new' couples and 17% of 'old' couples. Table 1 shows that the observed number was 12% engaged, 32% 'new' and 24% 'old'. But there was a social class bias in both ascertainment and response to study, so we must consider whether social class or any of the other variables might have influenced the results.

Table 3 summarizes the intercorrelations found. Age, religion, personality and father's occupation occasionally showed statistical significance but are clearly of little practical importance. We shall ignore them to concentrate on the more important factors. Education and social class are consistently correlated with smoking. Note that the sign of the correlation reflects the method of coding. A low social class (V) is represented by a higher number (5), while a poor education is represented by a low number. Therefore education is negatively, and class positively, correlated with smoking. This could mean that well educated/higher social class people are more often non-smokers, or that if they smoke, they smoke less. Table 4 shows the amount smoked by marriage cohort, sex, class, and education. The most obvious difference shown is that well educated/higher social class people are much more likely to be non-smokers, but there is also a tendency for the smokers among them to smoke less.

Table 4. Amount smoked by cohort, class, sex and education.

Group	Total	Nil	$\frac{1}{2}$–9	10–19	20–29	30–39	40+
(a) Engaged							
All men	68	51	2	4	11		
All women	68	54	7	5	2		
Men, Class I–IIIN	58	46	2	1	9		
Men, Class IIIM–V	10	5	0	3	2		
Women, Class I–IIIN	65	52	7	5	1		
Women, Class IIIM–V	3	2	0	0	1		
Men, Educ $\geqslant$ Highers	54	42	2	2	8		
Men, Educ $\leqslant$ 0 level	14	9	0	2	3		
Women, Educ $\geqslant$ Highers	56	48	4	4	0		
Women, Educ $\leqslant$ 0 level	12	6	3	1	2		
(b) New							
All men	112	51	16	16	20	5	4
All women	112	60	11	19	18	2	2
Men, Class I–IIIN	67	38	9	8	9	2	1
Men, Class IIIM–V	45	13	7	8	11	3	3
Women, Class I–IIIN	95	54	8	15	14	2	2
Women, Class IIIM–V	17	6	3	4	4	0	0
Men, Educ $\geqslant$ Highers	53	34	6	6	5	1	1
Men, Educ $\leqslant$ 0 level	59	17	10	10	15	4	3
Women, Educ $\geqslant$ Highers	58	38	7	4	6	1	2
Women, Educ $\leqslant$ 0 level	54	22	4	15	12	1	0
(c) Old							
All men	223	107	25	26	43	15	7
All women	223	150	20	23	27	2	1
Men, Class I–IIIN	131	77	17	15	19	2	1
Men, Class IIIM–V	92	30	8	11	24	13	6
Women, Class I–IIIN	174	122	13	19	19	0	1
Women, Class IIIM–V	49	28	7	4	8	2	0
Men, Educ $\geqslant$ Highers	78	48	9	10	10	1	0
Men, Educ $\leqslant$ 0 level	145	59	16	16	33	14	7
Women, Educ $\geqslant$ Highers	79	58	10	6	5	0	0
Women, Educ $\leqslant$ 0 level	144	92	10	17	22	2	1

Table 5. Correlation for smoking in different groups in the sample

	(A) by class				(B) by education			
	I+II+IIIN		IIIM+IV+V		With 'Highers' or degree		With O-levels or nothing	
	N	r SD	N	r SD	N	r SD	N	r SD
Engaged	58	0·543 ± 0·112****	10	−0·017 ± 0·354	54	0·651 ± 0·105****	14	−0·031 ± 0·289
New	67	0·348 ± 0·116**	45	0·202 ± 0·149	53	0·159 ± 0·138	59	0·300 ± 0·126
Old	131	0·284 ± 0·084***	92	0·256 ± 0·102*	78	0·280 ± 0·110*	145	0·286 ± 0·080***

In general, the women's smoking habits were correlated principally with their education, while in the case of the men, education and occupation were of similar importance. Is it possible that correlation for smoking is merely an artefact of assortative marriage for education and social class? And might the correlation vary in different classes?

Table 5 shows the results when the sample was partitioned, firstly by husband's social class (I, II and IIIN versus IIIM, IV and V), and secondly by husband's education (those with Highers or A-levels or the equivalent, versus those without). Ideally we should partition into individual social classes, but the numbers in each class are too small for that. The significance levels in each subgroup were reduced and the standard deviations were greatly increased. However, the values of r in each subgroup remained broadly similar, except for the near-zero correlation in the lower social class, low education subgroups of the engaged couples. This involved only a handful of couples so it seems reasonable to ignore them. On the basis of these figures, there is no reason to suppose that assortative marriage for smoking is greatly different in individual social classes.

Finally, table 6 demonstrates that the correlation for smoking remained high even when the effect of social class and education had been taken into account. This has been calculated by the method of partial correlations (see Snedecor and Cochran 1967). The converse is also true—the correlation for smoking was insufficient to account for those for social class and education. An unexpected finding was that the correlation for 'best job' might have been entirely secondary to that for education.

Table 6. Correlation for trait A which remains, when trait B is held constant

		Engaged	New	Old
Trait A	Trait B	r SD	r SD	r SD
Smoking	Best job	$0{\cdot}353 \pm 0{\cdot}117$**	$0{\cdot}242 \pm 0{\cdot}093$**	$0{\cdot}275 \pm 0{\cdot}065$****
Best job	Smoking	$0{\cdot}244 \pm 0{\cdot}121$*	$0{\cdot}483 \pm 0{\cdot}084$****	$0{\cdot}470 \pm 0{\cdot}060$****
Smoking	Education	$0{\cdot}367 \pm 0{\cdot}116$**	$0{\cdot}232 \pm 0{\cdot}094$*	$0{\cdot}281 \pm 0{\cdot}065$****
Education	Smoking	$0{\cdot}418 \pm 0{\cdot}114$***	$0{\cdot}527 \pm 0{\cdot}082$****	$0{\cdot}620 \pm 0{\cdot}053$****
Education	Best job	$0{\cdot}327 \pm 0{\cdot}118$**	$0{\cdot}263 \pm 0{\cdot}093$***	$0{\cdot}280 \pm 0{\cdot}065$****
Best job	Education	$0{\cdot}204 \pm 0{\cdot}122$	$0{\cdot}199 \pm 0{\cdot}094$*	$0{\cdot}062 \pm 0{\cdot}067$

4. Discussion

We have been using parametric methods throughout, i.e. assuming a Normal distribution. Although this is a very convenient assumption, it is hardly realistic to treat, say, religious practice, as Normally distributed. However, the calculation of r is very robust regarding departures from Normality. The standard deviation of r and the method of partial correlations are both rather more sensitive, but in the areas of interest the values of r found are quite clear-cut. The fairly large sample size also minimizes the chance of errors.

Smokers may lie to doctors about their consumption (Kozlowski *et al.* 1980). However, the strongest spur to deceit is in the anti-smoking clinic or GP's surgery where they are anxious to avoid unfavourable medical comment. No such source of 'dissonance' occurs in the present study, and the smoker's spouse made an extra witness.

The high correlation for smoking is found even in those who have not yet married, showing that smoking may be an important factor in selection of a partner. This is hardly surprising, as the offer of a cigarette is a standard opening gambit in the courtship ritual. Engaged partners are known to be similar in several ways (Burgess and Wallin 1944), and most theories of mate selection assign a key role to social class matching (Eckland 1968). Although smoking habits are interlinked with the other psychosocial factors measured, tables 3 and 6 suggest that they are nevertheless assorting independently with a correlation coefficient r in the range 0·25–0·35. Of course, there is always the possibility that this is merely mirroring some other factor not examined such as IQ. But if IQ was the underlying cause this should have been revealed in the correction for education.

One distinction between smoking and other psychological traits may be noted: the effect on a dissimilar marriage. If partners have dissimilar personalities or educational backgrounds, the tensions may become acute only after some years. However, once one has chosen a partner with unlike smoking habits, there seems no reason to expect this to cause strife later.

Some correlations vary with social class. For example, Harrison *et al.* (1976) found assortative marriage for IQ only in social classes I, II, and III and for height only in classes IV and V—but both correlations were strong enough to be significant in the overall sample. Table 5 shows that this was not the case with smoking. If there had been marked differences, the method of partial correlations used to compile table 6 would have been invalid.

The smoking habits of the parents of these couples is not known, and their own children are very young. Therefore it cannot be shown whether assortative marriage has influenced the frequency of the trait. Since smoking habits are influenced partly by parental example, this influence must be increased if marriages are concordant, as they tend to be.

The present data do not allow us to conclude whether couples are converging after marriage. The only sure test of this would be to revisit the entire sample after some time has elapsed, and remeasure all the variables. In fact, these repeat visits are now being made, and the results will be presented at a later date.

Acknowledgements

I am grateful to Professor A. E. H. Emery for his help and encouragement, to the Carnegie Trust for financial support, and to the Eugenics Society for assistance with research costs.

References

Burgess, E. W., and Wallin, P., 1944, Homogamy in personality charateristics. *Journal of Abnormal and Social Psychology*, **39**, 475–481.

Coleman, D. A., 1977, Assortative mating in Britain. In *Equalities and Inequalities in Family Life*, edited by R. Chester and J. Peel (London: Academic Press).

Eckland, B. K., 1968, Theories of mate selection. *Eugenics Quarterly*, **15**, (2), 71–84.

Eysenck, H. J., 1960, *The Structure of Human Personality*, 2nd edition (London: Methuen), pp. 44–50.

Harrison, G. A., Gibson, G. B., and Hiorns, R. W., 1976, Assortative mating for psychometric, personality and anthropometric variation in a group of Oxfordshire villages. *Journal of Biosocial Science*, **8**, 145–153.

Kozlowski, L. T., Herman, C. P., and Frecker, R. C., 1980, What researchers make of what cigarette smokers say: Filtering smokers' hot air. *Lancet*, **i**, 699–700.

Roberts, D. F., 1977, Assortative mating in man: Husband/wife correlations in physical characteristics. *Eugenics Society Bulletin*, Suppl. 2, 1–45.

Snedecor, G. W., and Cochran, W. G., 1967, *Statistical Methods*, 6th edition (Iowa State).

Spuhler, J. N., 1968, Assortative mating with respect to physical characteristics. *Eugenics Quarterly*, **15**, (2), 128–140.

Vandenberg, S. G., 1972, Assortative mating, or who marries whom? *Behavior Genetics*, **2**, (2/3), 127–157.

Address correspondence to: Dr G. C. Sutton, University Department of Human Genetics, Western General Hospital, Edinburgh EH4 2XU, Scotland.

Zusammenfassung. Partnerwahl nach Rauchgewohnheiten wurde bei 68 Verlobten, 112 frisch verheirateten und 223 seit sechs oder mehr Jahren verheirateten Paaren untersucht. Alle drei Gruppen zeigten eine recht starke Partnerkorrelation für Rauchen, $r=0{,}332 \pm 0{,}047$ ($\bar{x} \pm s$). Die Rauchgewohnheiten waren hochkorreliert mit sozialer Klasse und Erziehung, aber die Korrelation für das Rauchen blieb hochsignifikant, selbst wenn auf diese Faktoren hin korrigiert wurde. Da diese Korrelation auch bei denjenigen vorhanden war, die noch nicht geheiratet hatten, könnten die Rauchgewohnheiten ein wichtiger Faktor bei der Auswahl des Heiratspartners darstellen.

Résumé. L'assortiment matrimonial pour le tabagisme a été étudié dans 68 couples de fiancés, 112 couples nouvellement mariés et 223 couples mariés depuis six ans ou plus. Les trois groupes montraient tous une corrélation assez forte entre les membres de couples pour le tabagisme, $r=0{,}332 \pm 0{,}047$ (moyenne + écart-type). Le tabagisme était hautement associé à la classe sociale et à l'éducation, mais la corrélation pour le tabagisme restait hautement significative après correction pour ces facteurs. Vu que cette corrélation était présente même chez ceux qui n'étaient pas encore mariés, le tabagisme pourrait bien être un facteur important du choix du conjoint.

```
***EMAS 2980 EMAS*** E'          .   lloway                   WESTERN GENERAL
***EMAS 2980 EMAS*** EQH303 S.M.Holloway                   WESTERN GENERAL
***EMAS 2980 EMAS*** EQH303 S.M.Holloway                   WESTERN GENERAL
***EMAS 2980 EMAS*** EQH303 S.M.Holloway                   WESTERN GENERAL
***EMAS 2980 EMAS*** EQH303 S.M.Holloway                   WESTERN GENERAL
***EMAS 2980 EMAS*** EQH303 S.M.Holloway                   WESTERN GENERAL
***EMAS 2980 EMAS*** EQH303 S.M.Holloway                   WESTERN GENERAL
***EMAS 2980 EMAS*** EQH303 S.M.Holloway                   WESTERN GENERAL

050133 DART 25K LISTED T40 LP40

FILE 'DART'

%BEGIN;! ASSORTATIVE MATING -PROGRAM MK 11
%INTEGER F,ST,DRO,Q,I,K,L,ONE,J,CORLS,PMAX,CMAX,PNO,FN,STARK,FATE,SIGNIF
%REAL STEP,X,V,W,Z,NEU,EXO,XXO,QUE,TIME,R,SD,YY1,YY2,SIGTX,SIGTY
%STRING(15) S,T
%STRING(7)%ARRAY STREAM (1:3)
%REALARRAY MATRIX(1:410,1:200),RO,SDRO(51:131),Y(1:410,1:2),%C
B(0:9),PSY(1:12),DSTEP(40:43),RRO(1:4,1:4),SOC(0:6),MRANK(0:1,0:14)
%BYTEINTEGERARRAY COMM(1:127,0:7),FLAG,F,G(0:410),CO(-1:14,1:2),WD(0:7)
%INTEGERARRAY NRO,SIGRO(51:131),ID(0:7),TALLY(0:50),YY(1:4),OF(0:6),%C
XTAB(-1:14,-1:14),DBS(0:2,0:2)
%EXTERNALROUTINESPEC PROMPT (%STRING(15)T)
%EXTERNALROUTINESPEC DEFINE(%STRING(63) T)

%ROUTINE STAR;! SEPARATE SECTIONS
NEWLINE;%CYCLE J=51,1,131;RO(J)=-950;%REPEAT;
%IF STARK=0 %THENRETURN;PRINTSTRING('
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
');NEWLINES(2);STARK=0;%END

%ROUTINE LIST ;! LIST COUPLES ANALYSED;
I=2;PRINTSTRING('
LIST : ');%CYCLE P=1,1,PMAX;%IF FLAG(P)=0 %THEN ->L1
PRINT(MATRIX(P,1),3,0);I=I+1;%IF I=15 %THENSTART;NEWLINE;I=0;%FINISH;
L1: %REPEAT;NEWLINE;%END

%ROUTINE STEPSET ;! SET SCALES FOR GRAPHS , ETC
->L1 %UNLESS 39<FN<44;STEP=DSTEP(FN);%IF STEP>0.01 %THENRETURN
PROMPT('STEP = ');READ(STEP);DSTEP(FN)=STEP;%RETURN
L1: STEP=1;%IF 89<FN<101 %THEN STEP=0.1;%IF 58<FN<84 %THEN STEP=10;%END

%ROUTINE SAY EH;! FLAG MISTAKES
PRINTSTRING(' EH?     ');NEWLINE;PROMPT(' EH?     ');
L1: READSYMBOL(I);%IF I#10 %THEN ->L1;%END

%ROUTINE BUMF;! SKIP RUBBISH IN DATA
L1: READSYMBOL(I);%IF I='+'%OR I='*' %THENRETURN;->L1;%END

%REALFN DATE ;! CONVERT DATES TO DECIMAL
L1: READSYMBOL(I);->L1 %UNLESS '0'-1<I<'4';READSYMBOL(J);K=0
TIME=(J-1-'0'+ 10*(I-'0'))/365;READSYMBOL(I);READSYMBOL(J)
%IF I='9' %AND J='9' %THEN K=1;%IF '2'<J<'8' %THEN TIME=TIME-0.00274;
TIME=TIME+(J-1-'0'+10*(I-'0'))/12;READSYMBOL(I);READSYMBOL(J);
%IF K=1 %THENRESULT= -990;TIME=TIME+(J-'0'+10*(I-'0'));
%IF TIME>85 %THEN TIME=TIME-100;%RESULT=TIME;%END
```

```
%ROUTINE PSYCHO (%INTEGER L) ;! ANALYSE QUESTIONNAIRE
%CYCLE K=L,1,L+3;XXO=0;NEU=0;EXO=0;QUE=0;
L1: READSYMBOL(Q);%IF Q<'A'%THEN->L1
%CYCLE I=1,1,12;V=0;%IF Q='X' %THENSTART;XXO=XXO+1;Q='Q';%FINISH
%IF Q='Q'%THENSTART;QUE=QUE+1;V=0.5;%FINISH
%IF Q='Y' %THEN V=1;PSY(I)=V;EXO=EXO+V;READSYMBOL(Q);%REPEAT
%CYCLE I=2,4,10;NEU=NEU+PSY(I);NEU=NEU+PSY(I+1);%REPEAT
%IF XXO>2 %THENSTART; NEU=-950;EXO=-1900;QUE=-950;%FINISH
MATRIX(P,K)=NEU;MATRIX(P,K+4)=EXO-NEU;MATRIX(P,K+8)=QUE;%REPEAT;%END

%ROUTINE SOCIAL (%INTEGER Q) ;! ANALYSE SOCIAL CLASS
L1: READSYMBOL(J);%IF J<'A' %THEN->L1;%CYCLE I=0,1,1;V=-950
%CYCLE K=Q,2,Q+6;J='Q'-J;R=SOC(J);MATRIX(P,K+I)=R;
%IF V<0 %AND R>0 %THEN V=R;READSYMBOL(J);%REPEAT;
MATRIX(P,Q+8+I)=V;%REPEAT;
%CYCLE Q=31,1,34;READ(MATRIX(P,Q));%REPEAT;
%CYCLE Q=0,1,1;READ(J);K=0;L=0;%IF J>3 %THENSTART;
K=(J&48)>>4;L=(J&12)>>2;J=(J&3);%FINISH;MATRIX(P,157+Q)=J;
MATRIX(P,169+Q)=K;MATRIX(P,171+Q)=L;%REPEAT;READ(MATRIX(P,159))
READ(MATRIX(P,160));READ(J);MATRIX(P,14)=J;READ(MATRIX(P,19));
READ(MATRIX(P,161));READ(MATRIX(P,162));%CYCLE I=0,1,1;V=MATRIX(P,159+I)
%IF MATRIX(P,157+I)>1.2 %THEN V=V/2;MATRIX(P,163+I)=V+J+MATRIX(P,161+I)
%REPEAT;%END

%ROUTINE INSKF (%INTEGER L) ;! READ SKINFOLD
%CYCLE I=0,1,1;Z=0;%CYCLE J=L,2,L+6;READ(W);X=W+0.0001
%IF X<0 %THENSTART;X=-950;->L1;%FINISH;K=191+INTPT(X*10)-INTPT(X)*10;
%IF X<20 %THEN MATRIX(P,K)=MATRIX(P,K)+1;X=LOG(W)
L1: Z=Z+X;MATRIX(P,J+I)=X;%REPEAT;
%IF Z>0 %THEN Z=LOG(Z);MATRIX(P,L+8+I)=Z;%REPEAT;%END

%ROUTINE CARDIO (%INTEGER L) ;! PULSE & BP
%CYCLE I=0,1,1;READ(X);%IF X<1 %THEN X=-950;MATRIX(P,L+I)=X;READ(Z);
%IF Z<1 %THEN Z=-950;MATRIX(P,L+2+I)=Z;W=-950
%IF X>0 %AND Z>0 %THEN W=100*Z/X;MATRIX(P,L+4+I)=W;%REPEAT
%CYCLE I=0,1,1;READ(X);%IF X<1 %THEN X=-950;MATRIX(P,L+6+I)=X;
SKIPSYMBOL;READ(X);%IF X<1 %THEN X=-950;MATRIX(P,L+8+I)=X;%REPEAT;%END

%ROUTINE WEIGHTY (%INTEGER L) ;! INPUTS WEIGHTS
%CYCLE I=0,1,1;Z=-950;%CYCLE K=L,2,L+2;READ(V);READ(X);X=V*14+X
%IF X<0 %THEN X=-950;MATRIX(P,K+I)=X;%IF Z<0 %AND X>0 %THEN Z=X;%REPEAT
MATRIX(P,L+4+I)=Z;%REPEAT;%END

%ROUTINE LOFTY (%INTEGER L,K) ;! INPUT HEIGHT
%CYCLE Q=L,1,L+1;Z=-950;%CYCLE I=0,2,K;READ(V);READ(X);X=V*12+X
%IF X<0 %THEN X=-950;MATRIX(P,Q+I)=X;%IF Z<0%AND X>0 %THEN Z=X;%REPEAT
%IF K=2 %THENSTART; MATRIX(P,Q+4)=Z;->L1;%FINISH;%IF X<0 %THEN->L1;
Z=MATRIX(P,Q+2);%IF Z>0 %THEN X=(X+Z)/2;MATRIX(P,Q+6)=X;
L1: %REPEAT;%END

%ROUTINE ZERO2 ;! NO FOLLOWUP DATA AVAILABLE
MATRIX(P,9)=-950;MATRIX(P,15)=-950;MATRIX(P,20)=-950;MATRIX(P,173)=-950
MATRIX(P,174)=-950;MATRIX(P,39)=-950;MATRIX(P,40)=-950
MATRIX(P,87)=-950;MATRIX(P,88)=-950;%CYCLE Q=1,1,10;MATRIX(P,Q+50)=-950;
MATRIX(P,Q+76)=-950;MATRIX(P,Q+110)=-950;MATRIX(P,Q+132)=-950;%REPEAT
%CYCLE Q=1,1,4;MATRIX(P,142+Q)=-950;MATRIX(P,164+Q)=-950;%REPEAT;%END
```

```
%ROUTINE INPAIR ;! INPUT DATA ON ONE COUPLE
BUMF;READSYMBOL(I);READSYMBOL(J);MATRIX(P,6)=I;MATRIX(P,7)=J;
READ(MATRIX(P,1));READ(FATE);%IF FATE<2 %THEN ZERO2
MATRIX(P,8)=FATE;W=DATE;MATRIX(P,2)=W;Z=DATE;MATRIX(P,3)=Z;
X=DATE;MATRIX(P,4)=X;READ(MATRIX(P,5));%IF X<0 %THEN X=Z
SOCIAL(147);%CYCLE L=0,1,1;Z=DATE;MATRIX(P,21+L)=X-Z
MATRIX(P,35+L)=W-Z;%REPEAT;TIME=X
%CYCLE K=0,1,1;W=0;%CYCLE I=23,2,27;READ(V);%IF V<0%THEN V=-950;W=W+V;
MATRIX(P,I+K)=V;%REPEAT;MATRIX(P,29+K)=W;%REPEAT
READ(MATRIX(P,37));READ(MATRIX(P,38));CARDIO(41);
%CYCLE I=41,1,50;MATRIX(P,I+20)=MATRIX(P,I);%REPEAT;
LOFTY(89,2);WEIGHTY(71)
%CYCLE I=191,1,200;MATRIX(P,I)=0;%REPEAT;INSKF(101);
MATRIX(P,99)=MATRIX(P,109);MATRIX(P,100)=MATRIX(P,110);PSYCHO(121);
%IF FATE<2 %THEN->L3;ONE=ONE+1;X=TIME;MATRIX(P,9)=X-DATE;I=0;
L1: READSYMBOL(J);%IF J<'A' %THEN->L1;J='Q'-J;R=SOC(J);
%IF J>0 %THEN MATRIX(P,153+I)=R;MATRIX(P,145+I)=R;
I=I+1;%IF I=1 %THEN ->L1;READ(MATRIX(P,39));READ(MATRIX(P,40))
CARDIO(51);READ(X);MATRIX(P,13)=X;READ(MATRIX(P,20));
%CYCLE I=0,1,1;READ(Z);MATRIX(P,165+I)=Z;
V=MATRIX(P,159+I);%IF MATRIX(P,157+I)>1.2 %THEN V=V/2
MATRIX(P,167+I)=Z+X+V;%REPEAT;
%CYCLE I=51,1,60;X=MATRIX(P,I+10);Z=MATRIX(P,I);%IF Z<0 %THEN->L2
%IF X>0 %AND X<Z %THEN->L2;MATRIX(P,I+10)=Z;
L2: %REPEAT;%CYCLE I=0,1,1;
MATRIX(P,65+I)=MATRIX(P,63+I)*100/MATRIX(P,61+I);%REPEAT
LOFTY(87,0);WEIGHTY(77);INSKF(111);
%CYCLE I=0,1,1;X=MATRIX(P,119+I);%IF X>0 %THENSTART;Z=MATRIX(P,99+I);
%IF Z>0 %THEN X=(X+Z)/2;MATRIX(P,99+I)=X;%FINISH;%REPEAT
READ(MATRIX(P,173));READ(MATRIX(P,174));PSYCHO(133);
L3: %CYCLE I=0,1,1;X=MATRIX(P,93+I);Z=-950
MATRIX(P,83+I)=(4.4*I)+(MATRIX(P,75+I)*3124.4/(X*X))
%IF FATE>4 %THEN Z=(4.4*I)+(MATRIX(P,81+I)*3124.4/(X*X))
MATRIX(P,85+I)=Z;%REPEAT;BUMF;%END

%ROUTINE ALLCOMS ;! READ IN SET OF VALID COMMANDS
SOC(0)=-950;SOC(1)=1;SOC(2)=2;SOC(3)=3;SOC(4)=3.5;SOC(5)=4
SOC(6)=5;%CYCLE CMAX=1,1,30;COMM(CMAX,0)=0;%REPEAT;
%CYCLE CMAX=31,1,50;COMM(CMAX,0)=CMAX-30;%REPEAT;I=21
%CYCLE CMAX=51,1,127;COMM(CMAX,0)=I;I=I+2;%REPEAT;CMAX=0
L1: READSYMBOL(J);%IF J='.' %THENRETURN;
%IF J<'A'%THEN->L1;CMAX=CMAX+1;I=0
L2: I=I+1;COMM(CMAX,I)=J;READSYMBOL(J);%IF J>='0'%THEN->L2
L3: %IF I=7 %THEN->L1;I=I+1;COMM(CMAX,I)=' ';->L3;%END

%INTEGERFN RWORD ;! READ + IDENTIFY A COMMAND WORD
WD(0)=0;%CYCLE J=1,1,7;WD(J)=' ';%REPEAT;J=1;
L1: READSYMBOL(I);%IF I<'A'%THEN->L1
L2: WD(J)=I;READSYMBOL(I);%IF I<'0'%THENSTART;I=0;->L3;%FINISH;
%IF J<7 %THEN J=J+1;->L2
L3: I=I+1;%IF I>CMAX%THEN %RESULT=0;J=1
L4: %IF COMM(I,J)#WD(J)%THEN->L3
J=J+1;%IF J<8%THEN->L4;WD(0)=COMM(I,0);%RESULT=I;%END

%ROUTINE ALLPAIRS(%INTEGER MODE);! READ IN ALL COUPLES
SELECTINPUT(ST);T=STREAM(ST);P=PNO;ONE=0;%IF MODE>1 %THEN ->L1;
S='COMPLETE';ALLCOMS;PMAX=410;CORLS=0;
L1: READSYMBOL(Q);%IF Q#'*'%THEN->L1;%IF NEXTSYMBOL='!'%THEN->L2
%IF P=PMAX %THENSTART;S='INCOMPLETE';->L2;%FINISH;P=P+1;INPAIR;->L1
L2: PRINTSTRING('INPUT FROM FILE '.T.' IS '.S.' WITH');WRITE((P-PNO),3)
PRINTSTRING(' PAIRS (');WRITE(ONE,3);PRINTSTRING(' REVISITED)
');CORLS=CORLS+ONE;PNO=P;SELECTINPUT(0);CLOSESTREAM(ST);NEWLINE;
%IF 0<MODE<3 %THENRETURN;%CYCLE P=1,1,410;Y(P,1)=0;Y(P,2)=0;
```

```
%ROUTINE TAG ?; ECHO INPUT COMMAND, IF CORRECT
%IF 12> ORD>8 %THENRETURN;J=1;NEWLINE;
%IF ORD=4%AND 89<ID(J+1)<101%THENSTART
PRINTSTRING('THEIR TOTSKF ');%RETURN;%FINISH;
L1: L=ID(J);%CYCLE K=1,1,6;PRINTSYMBOL(COMM(L,K));%REPEAT
L=ID(J+1);%CYCLE K=1,1,7;PRINTSYMBOL(COMM(L,K));%REPEAT;L=ID(4)
%IF J=5 %OR L=0 %THENRETURN;J=5;SPACE;PRINTSYMBOL(L);SPACE;->L1;%END

%INTEGERFN IDENT (%INTEGER MODE) ;! IDENT INPUT COMMAND
K=1;%CYCLE L=0,1,7;ID(L)=0;%REPEAT
L1: PROMPT('WHO : ');L=RWORD;%IF L<27%OR L>29%THEN->L1;
L2: PROMPT('TRAIT : ');Q=RWORD;%IF Q<31%THEN->L2
%IF Q<49%AND L<29%THEN ID(0)=1;%IF Q>48 %AND L=29 %THEN ID(0)=1;
%IF 48<Q<51 %AND L=27 %THEN ID(0)=1
J=WD(0)+L-27;%IF L=29 %THEN J=J-2;ID(K)=L;ID(K+1)=Q;ID(K+2)=J;
%IF K>1 %OR MODE=1 %THEN->L4;ID(4)='V';%IF ORD#8 %THEN ->L3
L6: PROMPT('OF : ');READSYMBOL(K);
%CYCLE I=1,1,4;%IF K=OF(I)%THEN->L7;%REPEAT;->L6
L7: ID(4)=K;OF(0)=I;K=5;->L1
L3: PROMPT('V : ');K=RWORD;%IF K#30%THEN ->L3;K=5;->L1;
L4: FN=Q;%IF ORD=8 %THEN ->L5;%IF FN<39 %OR FN>43 %THENSTART;
%IF MODE=1 %THEN->L5;FN=ID(2);->L5 %UNLESS 44>FN>39;%FINISH;I='0'+FN-39
%IF DSTEP(FN)>-4 %THEN ->L5;NEWLINE;PRINTSTRING('THEIR TRAIT');
PRINTSYMBOL(I);PRINTSTRING('  NOT YET DEFINED');NEWLINE;%RESULT=1
L5: %IF ID(3)=ID(7) %THEN ID(0)=1;%IF ID(0)=0 %THEN TAG %ELSE SAY EH;
%RESULT=ID(0);%END

%ROUTINE PAIRS (%INTEGER MODE);! PUT DATA IN ARRAY Y FOR EASY REF
ONE=1;K=ID(7);J=ID(3)
%IF J#YY(1)%THEN->L1;%IF MODE=1 %OR K=YY(2)%THEN->L4;L=2;->L3
L1: %IF J#YY(2)%THEN->L2;ONE=2;%IF MODE=1 %OR K=YY(1) %THEN->L4;L=1;->L3
L2: %IF YY(2)=0 %OR K=YY(1)%THEN ONE=2;%IF YY(1)=0 %THEN ONE=1;
YY(ONE)=J;%CYCLE Q=1,1,PMAX;Y(Q,ONE)=MATRIX(Q,J);%REPEAT
L=3-ONE;%IF MODE=1 %OR K=YY(L) %THEN->L4
L3: YY(L)=K;%CYCLE Q=1,1,PMAX;Y(Q,L)=MATRIX(Q,K);%REPEAT
L4: %CYCLE P=1,1,PMAX;%IF FLAG(P)=0 %THENSTART;Y(P,1)=-990;
Y(P,2)=-990;%FINISH;%REPEAT;J=0;K=0;%CYCLE Q=1,1,PMAX;
%IF Y(Q,1)>-900 %THEN J=J+1;%IF Y(Q,2)>-900 %THEN K=K+1;%REPEAT
YY(3)=J;YY(4)=K;%END

%ROUTINE SIGDER (%INTEGER N,I);! SD + SIGNIF LEVEL
SIGNIF=0;CORLS=CORLS+1;%IF I=1 %THENSTART;X=R;->L1;%FINISH
SD=1;%IF N<10 %THENRETURN;SD=SQRT((1-R*R)/(N-2));X=R/SD;
L1: %IF X<0 %THEN X=-X
%IF X<1.96 %THENRETURN;SIGNIF=1;%IF X>2.58 %THEN SIGNIF=2
%IF X>3.29 %THEN SIGNIF =3;%IF X>3.89 %THEN SIGNIF=4;%END

%ROUTINE TWOBYTWO ;!  2 * 2 CONTINGENCY TABLE
V=10;%CYCLE I=1,1,2;%CYCLE J=1,1,2;W=(OBS(I,0)*OBS(0,J)/PNO);
%IF W<V %THEN V=W;%REPEAT;%REPEAT;PRINTSTRING('
CHI-SQUARED :                      ');%IF V>4.99 %THEN ->L2;
%IF V<1.01 %AND PNO>39 %THEN ->L2;
L1: PRINTSTRING('IS NOT COMPUTABLE');%RETURN
L2: X=OBS(1,1)*OBS(2,2)-OBS(1,2)*OBS(2,1);%IF X<0 %THEN X=-X;
R=(X-PNO/2)**2 *PNO;X=OBS(0,1)*OBS(0,2)*OBS(1,0)*OBS(2,0)
%IF X<0.00001 %THEN ->L1;R=R/X;PRINTSTRING('DF= 1    X2=')
PRINT(R,3,3);CORLS=CORLS+1;%END
```

```
%ROUTINE SPLIT;! ALL OR NONE TRAIT
%INTEGER M,F,A;
%IF IDENT(2)=1 %THENRETURN;PAIRS(2);%CYCLE I=0,1,2;%CYCLE J=0,1,2;
OBS(I,J)=0;%REPEAT;%REPEAT;PRINTSTRING(' SPLIT ');R=-950;
PROMPT('CUT-OFF = ');READ(X);READ(Z);PNO=0;
%CYCLE Q=1,1,PMAX;V=Y(Q,1);W=Y(Q,2);%IF V<-900 %OR W<-900 %THEN->L1;
I=1;J=1;%IF V>X %THEN I=2 ;%IF W>Z %THEN J=2 ;
OBS(I,J)=OBS(I,J)+1;PNO=PNO+1;
L1: %REPEAT;%IF ONE=2%THENSTART;K=OBS(1,2);OBS(1,2)=OBS(2,1);OBS(2,1)=K
%FINISH;PRINT(X,3,1);PRINT(Z,3,1);%IF PNO<10 %THENRETURN;PRINTSTRING('
CONCORDANCE :         0-0         0-1         1-0         1-1
NUMBER + %   : ');%CYCLE I=1,1,2;%CYCLE J=1,1,2;K=OBS(I,J);WRITE(K,3)
OBS(I,0)=OBS(I,0)+K;OBS(0,J)=OBS(0,J)+K;PRINTSYMBOL('(');
PRINT((K/PNO)*100,2,0);PRINTSTRING('%)');%REPEAT;%REPEAT;
TWOBYTWO;A=OBS(2,2);M=OBS(0,2);F=OBS(2,0)
V=(M-(M*M)/PNO);W=(F-(F*F)/PNO);R=SQRT(V*W);PRINTSTRING('
CORRELATION :                    ');%IF R>0.000001 %THEN ->L2;R=-950;%RETURN
L2: R=(A-(M*F)/PNO)/R;
%IF R>0.99 %OR R< -0.99 %THEN R=-950 %ELSE SIGGER(PNO,0);%END

%ROUTINE EVALRO ;! COMPUTE OR LOOK-UP R-SP
%REAL SY1Y2,SQY1,SQY2,SIGY1,SIGY2
%IF ORD # 11  %AND IDENT(2)=1 %THENRETURN;FN=0;SIGY1=0;SIGY2=0;SY1Y2=0;
SQY1=0;SQY2=0;PNO=0;L=ID(3)-ID(7);L=L*L;
%IF L>1 %OR ID(2)#ID(6) %THEN ->L10;FN=ID(2);->L10 %UNLESS 50<FN<132;
R=RO(FN);%IF R<-900 %THEN ->L10;
SD=SDRO(FN);PNO=NRO(FN);SIGNIF=SIGRO(FN);%RETURN
L10: PAIRS(2);P=0;X=5;%IF YY(3)<10 %OR YY(4)<10 %THEN ->L4
L1: P=P+1;%IF P>PMAX %THEN->L2;YY1=Y(P,1);YY2=Y(P,2);
%IF YY1<-900 %OR YY2<-900 %THEN->L1;PNO=PNO+1;SY1Y2=SY1Y2+(YY1*YY2)
SQY1=SQY1+(YY1*YY1);SQY2=SQY2+(YY2*YY2);SIGY1=SIGY1+YY1;
SIGY2=SIGY2+YY2;->L1
L2: %IF PNO<11 %THEN->L4;R=(SQY1-SIGY1*SIGY1/PNO)*(SQY2-SIGY2*SIGY2/PNO)
R=SQRT(R);%IF R<0.000001 %THEN->L4;R=(SY1Y2-SIGY1*SIGY2/PNO)/R
%IF R>0.99 %THEN ->L4;SIGGER(PNO,0);
%IF FN<51 %THENRETURN;RO(FN)=R;
SDRO(FN)=SD;NRO(FN)=PNO;SIGRO(FN)=SIGNIF;%RETURN
L4: %IF FN>50 %THEN RO(FN)=-950;R=-950;%END

%ROUTINE PRCORL ;! PRINT CORREL COEFF R-SP
T='    R=';%IF ORD=15 %THEN T='  RS=';
%IF R<-900 %THEN ->L1;PRINTSTRING('   N=');WRITE(PNO,3);
PRINTSTRING(T);PRINT(R,3,3);PRINTSTRING(' SD=');
PRINT(SD,1,3);%IF SIGNIF =0 %THEN->L2;SPACES(3);
%CYCLE J=1,1,SIGNIF;PRINTSYMBOL('*');%REPEAT;->L2
L1: PRINTSTRING('  IS NOT COMPUTABLE')
L2: NEWLINE;%END

%ROUTINE BIAS ;! DETECT MEASUREMENT BIAS
%IF IDENT(1)=1 %THENRETURN;PAIRS(1);
STEPSET;X=STEP/10;%CYCLE P=0,1,9;B(P)=0;%REPEAT;
P=0;->L1 %UNLESS 89<FN<101;X=0.1;%CYCLE P=1,1,PMAX;
%IF FLAG(P)=1 %THENSTART;%CYCLE Q=0,1,9;B(Q)=B(Q)+MATRIX(P,191+Q);
%REPEAT;%FINISH;%REPEAT;->L2
L1: P=P+1;%IF P>PMAX %THEN->L2;W=Y(P,ONE);%IF W<-900 %THEN->L1
Q=INTPT(FRACPT(W/STEP)*10+0.00001);B(Q)=B(Q)+1;->L1
L2: NEWLINE;PRINTSTRING('SCALE READING : ');
%CYCLE P=0,1,9;PRINT(X*P,1,2);%REPEAT;PRINTSTRING('
NO OF ENTRIES : ');%CYCLE P=0,1,9;PRINT(B(P),4,0);%REPEAT;NEWLINE;%END
```

```
%INTEGERFN REXPR ;! RECOGNISE LOGICAL EXPRESSION
%REAL LX,HX ; %INTEGER NO
NO=0;J=TALLY(0);
L1: PROMPT('SYMBOL :');READSYMBOL(I);%IF I=' '%THEN->L1
%IF I='\'%THENSTART;NO=1-NO;->L1;%FINISH;K=' ';%IF NO=1 %THEN K='\'
TALLY(J)=K+200;J=J+1;K=0;%IF I='='%THEN K=2;%IF I='<'%THEN K=1;
%IF I='>'%THEN K=3;%IF K=0 %THENRESULT=0;TALLY(J)=I+200;J=J+1
L4: PROMPT('NUMBER = ');%IF NEXTSYMBOL=' ' %THENSTART;SKIPSYMBOL;->L4;
%FINISH;%IF NEXTSYMBOL#''''%THENSTART;READ(X);->L2;%FINISH
SKIPSYMBOL;READSYMBOL(I);X=I;READSYMBOL(I);%IF I#''''%THENRESULT=0
L2: %IF X<-940 %THEN %RESULT=0;LX=X*0.9995;HX=X*1.0005
TALLY(J)=-INT((X+1000)*100);TALLY(0)=J+1;J=ID(3)
%CYCLE P=1,1,PMAX;L=0;W=MATRIX(P,J);%IF W<-900 %THEN->L3
I=2;%IF LX>W %THEN I=1;%IF HX<W%THEN I=3;
%IF I=K %THEN L=1;%IF NO=1 %THEN L=1-L
L3: G(P)=L;%REPEAT;%RESULT=1; %END

%ROUTINE RESTRICT ;! LOOK ONLY AT SUBSET OF DATA
%INTEGER CD ; %SWITCH SET(24:26)
%CYCLE Q=0,1,31;TALLY(Q)=0;%REPEAT;TALLY(1)=9
PROMPT('SELECT :');Q=RWORD;%IF Q>26 %OR Q<24 %THENSTART;SAY EH;%RETURN
%FINISH;TALLY(2)=Q;TALLY(0)=3;NEWLINE;->SET(Q)
SET(24): %CYCLE Q=1,1,PMAX;F(Q)=1;%REPEAT;->L19;
SET(25): %CYCLE Q=1,1,PMAX;F(Q)=FLAG(Q);%REPEAT;
L15: PROMPT('COND :');CD=RWORD;%IF CD>23 %OR CD <22 %THENSTART;SAY EH;
%RETURN;%FINISH;J=TALLY(0);TALLY(J)=CD;TALLY(0)=J+1;->L17
SET(26): %CYCLE Q=1,1,PMAX;F(Q)=0;%REPEAT;CD=22;
L17: %IF IDENT(1)=1 %THENRETURN;J=TALLY(0);TALLY(J)=ID(1);
TALLY(J+1)=ID(2);TALLY(0)=J+2;%IF REXPR=0 %THENSTART;SAYEH;
%RETURN;%FINISH;%CYCLE Q=1,1,PMAX;K=F(Q);L=G(Q);
%IF CD=23 %THEN F(Q)=K&L %ELSE F(Q)=L ! K;%REPEAT;
L18: Q=NEXTSYMBOL;%IF Q=10 %THEN->L19;%IF Q<'A' %THENSTART;
SKIPSYMBOL;->L18;%FINISH;->L15;
L19: L=0;K=0;%CYCLE P=1,1,PMAX;J=F(P);%IF J#FLAG(P)%THENSTART
L=1;FLAG(P)=J;%FINISH;K=K+J;%REPEAT;%CYCLE J=1,1,TALLY(0)-1
P=TALLY(J);%IF P<0 %THENSTART;PRINT((-P/100)-1000,1,2);SPACE;->L21;
%FINISH;%IF P>200 %THENSTART;PRINTSYMBOL(P-200);->L21;%FINISH
%CYCLE Q=1,1,7;I=COMM(P,Q);%IF I#' ' %THEN PRINTSYMBOL(I);%REPEAT;SPACE
L21: %REPEAT;PNO=K;%IF L=1 %THEN->L20;PRINTSTRING('
NO CHANGE MADE ');NEWLINE;%RETURN
L20: STAR;WRITE(K,1);PRINTSTRING(' COUPLES ANALYSED');ST=1;YY(1)=0
YY(2)=0;FN=0;%IF K<11 %THENSTART;PRINTSTRING('   SAMPLE INQUORATE');
ST=0;%FINISH;NEWLINE;%END

%ROUTINE CHANGE ;! ALTER DATABASE
%IF IDENT(1)=1 %THENRETURN;Q=ID(3);I=WD(0);%IF I>50 %THEN RD(I)=-950;
L1: PROMPT('PAIRNO = ');READ(P);%RETURN %UNLESS 0<P<PMAX
PROMPT('VALUE  = ');READ(X);MATRIX(P,Q)=X;K=0
%IF  Q=YY(1) %THEN K=1;%IF Q=YY(2) %THEN K=2
%IF K#0 %THEN Y(P,K)=X;->L1;%END

%ROUTINE CREATE ;! GENERATE NEW TRAIT ENTITY
%SWITCH LOOP (1:6)
%IF IDENT(2)=1 %THENRETURN;PROMPT('TRAIT NO = ');
L1: READ(I);->L1 %UNLESS 0<I<3;DSTEP(I+39)=-1;PAIRS(2);J=OP(0);
%IF J=2 %AND ONE=2 %THEN J=5;%IF J=4 %AND ONE=2 %THEN J=6
%CYCLE Q=1,1,PMAX;V=Y(Q,1);W=Y(Q,2);-> LOOP(J)
LOOP(1): X=V+W;->L2
LOOP(2): X=V-W;->L2;
LOOP(3): X=V*W;->L2
LOOP(4): %IF 0.0001>W>-0.0001 %THEN X=-950 %ELSE X=V/W;->L2
LOOP(5): X=W-V;->L2;
LOOP(6): %IF 0.0001>V>-0.0001 %THEN X=-950 %ELSE X=W/V
L2: %IF V<-900 %OR W< -900 %THEN X=-950;MATRIX(Q,9+I)=X;%REPEAT;
PRINTSTRING(' IS THEIR TRAIT');I='0'+I;PRINTSYMBOL(I);NEWLINE;%END
```

```
%ROUTINE FIND ;! FIND COUPLE WITH GIVEN MEASUREMENT
%IF IDENT(1)=1 %THENRETURN;PAIRS(1);
L1: PROMPT('TARGET = ');READ(X);%IF X<-900 %THEN %RETURN
PRINTSTRING(' TARGET ');PRINT(X,3,1);STEPSET;W=STEP/10;W=W*W;I=0
%CYCLE Q=1,1,PMAX;Z=Y(Q,ONE)-X;Z=Z*Z;%IF Z<W %THENSTART
I=1;NEWLINE;PRINT(MATRIX(Q,1),3,0);PRINT(Y(Q,ONE),3,2);%FINISH
%REPEAT;%IF I=0 %THEN PRINTSTRING(' NOT FOUND ');NEWLINE;->L1;%END

%ROUTINE GRAPH;! DRAW GRAPH OF DATA
%REAL M,MEAN,TOT,MAX,MIN
%IF IDENT(1)=1 %THENRETURN;PAIRS(1);%IF YY(3)<3 %THENSTART
PRINTSTRING(' NOT GRAPHABLE');%RETURN;%FINISH;MIN=1000;MAX=-900
STEPSET;TOT=0;%CYCLE P=0,1,50;TALLY(P)=0;%REPEAT;P=0;PNO=0
L2: P=P+1;%IF P>PMAX %THEN->L3;M=Y(P,ONE);%IF M<-900 %THEN->L2
PNO=PNO+1;%IF M<MIN %THEN MIN=M;
TOT=TOT+M;%IF M>MAX %THEN MAX=M;->L2
L3: NEWLINE;PRINTSTRING('N = ');WRITE(PNO,3);PRINTSTRING('  RANGE IS ')
PRINT(MIN,1,2);PRINTSTRING(' TO ');PRINT(MAX,1,2);
P=0;MEAN=TOT/PNO;PRINTSTRING('
MEAN =');PRINT(MEAN,2,2);TOT=0;MIN=INT(MIN/STEP)*STEP;
L4: P=P+1;%IF P>PMAX %THEN->L5;M=Y(P,ONE);%IF M<-900 %THEN->L4
I=1+INTPT((M-MIN)/STEP);%IF I<0%THEN I=0;
%IF I>50%THEN I=50;TALLY(I)=TALLY(I)+1;TOT=TOT+(M-MEAN)**2;->L4
L5: PRINTSTRING('      SD = ');PRINT(SQRT(TOT/(PNO-1)),2,2);NEWLINE;
L=1;%CYCLE K=49,-1,5;%IF TALLY(K)>0%THEN ->L9;%REPEAT;->L1
L9: L=L+1;%IF L=K-2 %THEN->L1;%IF TALLY(L)=0 %THEN->L9;%IF L<4 %THEN L=1
L1: %IF L>1 %THEN TALLY(0)=TALLY(0)+TALLY(1);%CYCLE P=L,1,K;NEWLINE;
PRINT((P-1)*STEP+MIN,3,1);PRINTSYMBOL(':');I=TALLY(P);
L6: %IF I>50 %THENSTART;PRINTSTRING('FIFTYPLUS!');I=I-50;->L6;%FINISH
L7: %IF I>9 %THENSTART;PRINTSTRING('+++++++++!');I=I-10;->L7;%FINISH
L8: %IF I>0 %THENSTART;PRINTSYMBOL('+');I=I-1;->L8;%FINISH;%REPEAT;
NEWLINE;I=TALLY(0);J=TALLY(50);%IF I+J=0 %THENRETURN
PRINTSTRING('PLUS ');WRITE(I,1);PRINTSTRING(' LESSER AND')
WRITE(J,1);PRINTSTRING(' GREATER VALUES
');%END


%REALFN SEMI (%INTEGER A,B,C) ; ! CORRECT A & B FOR C
R=(1-RRO(A,C)**2)*(1-RRO(B,C)**2);%IF R<0.0001 %THEN %RESULT=-950
%RESULT=(RRO(A,B)-RRO(A,C)*RRO(B,C))/SQRT(R);%END

%ROUTINE PARTCOR ; ! CANONICAL CORRELATIONS
%REAL PX,QX,RX;%INTEGER A,B,C,D,E,NTOT;%INTEGERARRAY TR(1:4)
%CYCLE A=1,1,4;RRO(A,A)=1;%REPEAT;
L1: PROMPT('TRAIT : ');A=RWORD;%IF A<31 %THEN ->L1;TR(1)=WD(0);
L2: PROMPT('V : ');B=RWORD;%IF  B#30 %THEN ->L2;TR(2)=TR(1)+1
L3: PROMPT('TRAIT : ');B=RWORD;%IF B<31 %THEN->L3;TR(3)=WD(0)
E=0;TR(4)=TR(3)+1;NTOT=PMAX;PRINTSTRING('
1)  CORREL MATRIX ');%CYCLE C=1,1,7;PRINTSYMBOL(COMM(A,C));%REPEAT;
PRINTSTRING(' V ');%CYCLE C=1,1,7;PRINTSYMBOL(COMM(B,C));%REPEAT
%CYCLE C=1,1,3;ID(3)=TR(C);ID(2)=A
%IF C=3 %THEN ID(2)=B;ID(1)=27;%IF C=2 %THEN ID(2)=28
%CYCLE D=C+1,1,4;ID(7)=TR(D);ID(6)=B;%IF D=2 %THEN ID(6)=A;ID(5)=28;
%IF D=3 %THEN ID(5)=27; EVALRO;%IF R<-900 %THEN E=1
RRO(C,D)=R;RRO(D,C)=R;%IF PNO<NTOT %THEN NTOT=PNO;%REPEAT;%REPEAT;
%IF NTOT<50 %THEN E=1;NEWLINE;
%CYCLE A=1,1,4;NEWLINE;%CYCLE B=1,1,4;R=RRO(A,B);
%IF R<-1 %THEN PRINTSTRING('      ..?..') %ELSE PRINT(R,5,3);%REPEAT;
%REPEAT;NEWLINES(2);PRINTSTRING('2) 1ST PARTCORRECTED FOR 2ND ');
%IF E=0 %THEN ->L4;PRINTSTRING(' NOT COMPUTABLE');NEWLINE;%RETURN;
L4: PX=SEMI(1,2,4);QX=SEMI(1,3,4);RX=SEMI(2,3,4)
R=(PX-QX*RX)/SQRT((1-QX*QX)*(1-RX*RX));SIGGER((NTOT-2),0)
PRINTSTRING('  N=');WRITE(NTOT,3);PRINTSTRING('   R=');PRINT(R,3,3);
PRINTSTRING(' SD=');PRINT(SD,1,3);%IF SIGNIF=0 %THEN->L5;
SPACES(3);%CYCLE A=1,1,SIGNIF;PRINTSYMBOL('*');%REPEAT;
L5: NEWLINE;%IF E=1 %THENRETURN;RRO(1,2)=RRO(4,3);RRO(3,4)=RRO(2,1);
E=1;RRO(2,1)=RRO(1,2);RRO(4,3)=RRO(3,4);RRO(1,4)=RRO(2,3)
RRO(2,3)=RRO(4,1);RRO(4,1)=RRO(1,4);RRO(3,2)=RRO(2,3);PRINTSTRING('
3) 2ND PARTCORRECTED FOR 1ST ');->L4;%END
```

```
%ROUTINE TTEST; ! T TEST FOR MATCHED PAIRS
%REAL DI,SIG2DI,SIGDI2,DMU
%IF IDENT(2)=1 %THEN %RETURN;PAIRS(2);%IF YY(3)<8 %OR YY(4)<8 %THEN ->L3
P=0;PND=0;SIGDI2=0;SIG2DI=0;
L1: P=P+1;%IF P>PMAX %THEN->L2;YY1=Y(P,1);YY2=Y(P,2);
%IF YY1<-900 %OR YY2 <-900 %THEN ->L1;PND=PND+1;DI=YY1-YY2;
SIG2DI=SIG2DI+DI;SIGDI2=SIGDI2+(DI*DI);->L1
L2: %IF PND<8 %THEN ->L3;DMU=SIG2DI/PND;SIG2DI=SIG2DI*SIG2DI
SD=SQRT((SIGDI2-(SIG2DI/PND))/(PND-1));R=DMU*SQRT(PND)/SD
PRINTSTRING(' DF=');WRITE(PND-1,3);PRINTSTRING('    T=');
PRINT(R,3,3);PRINTSTRING(' SD=');PRINT(SD,1,3);
SIGGER(PND-1,1);%IF SIGNIF=0 %THEN->L4;SPACES(3);
%CYCLE J=1,1,SIGNIF;PRINTSYMBOL('*');%REPEAT;->L4
L3: PRINTSTRING(' IS NOT COMPUTABLE')
L4: NEWLINE; %END

%ROUTINE ANOVA ;! ANALYSIS OF VARIANCE
L1: PROMPT('TRAIT : ');READSYMBOL(I);%IF I#10 %THEN ->L1
PRINTSTRING('SORRY!  ANALYSIS OF VARIANCE NOT YET IMPLEMENTED
');%END

%ROUTINE XTABLE ; ! CROSS TABULATE TRAITS
%INTEGER HI,HJ,LOI;%REAL V1,V2,V3,V4,V5,V6
%IF IDENT(2)=1 %THENRETURN;PROMPT('STEP = ');READ(V1);READ(V2);LOI=1;
READ(V3);READ(V4);READ(V5);READ(V6);PRINT(V1,5,1);PRINT(V2,1,1)
PRINT(V3,1,1);PRINT(V4,3,1);PRINT(V5,1,1);PRINT(V6,1,1)
%IF V2<0.001 %OR V5<0.001 %THEN->L1
%IF V3<V1 %OR V6<V4 %THEN->L1;HI=INTPT(0.001+(V3-V1)/V2)+2;
HJ=INTPT(0.001+(V6-V4)/V5)+2;%IF 5<HJ<15 %AND 5<HI<15 %THEN ->L2
L1: NEWLINE;ORD=0;PRINTSTRING('CROSS TABLE MALFORMED');NEWLINE;%RETURN
L2: %CYCLE I=2,1,14;CO(I,1)=' ';CO(I,2)=' ';%REPEAT;CO(HI-1,1)='V';
CO(HI-1,2)='3';CO(HI,1)='H';CO(HI,2)='I';PAIRS(2);
%CYCLE I=-1,1,14;%CYCLE J=-1,1,14;XTAB(I,J)=0;%REPEAT;%REPEAT;K=5;L=5;
%CYCLE P=1,1,PMAX;I=-1;J=-1;W=Y(P,1);X=Y(P,2);%IF W<-900 %THEN ->L3;I=0
%IF W<V1 %THEN->L3;I=HI;%IF W>V3 %THEN->L3;I=INTPT(0.001+(W-V1)/V2)+1
L3: %IF X<-900 %THEN ->L4;J=0;%IF X<V4 %THEN->L4;J=HJ;
%IF X>V6 %THEN ->L4;J=INTPT(0.001+(X-V4)/V5)+1
L4: XTAB(I,J)=XTAB(I,J)+1;%IF I>K %THEN K=I;%IF J>L %THEN L=J;
%IF LOI>I %THEN LOI=I;%REPEAT;%IF ORD=15 %THEN %RETURN;PRINTSTRING('
2ND VAR  ?? LO V4  -> ');J=(HJ*3)-16;%IF J>1 %THEN SPACES(J);
PRINTSTRING(' -> V6 HJ
1ST VAR');%CYCLE I=LOI,1,K;NEWLINE;PRINTSYMBOL(CO(I,1));
PRINTSYMBOL(CO(I,2));SPACES(6);
%CYCLE J=-1,1,L;K=XTAB(I,J);%IF K>0 %THEN WRITE(K,2) %ELSE SPACES(3);
%REPEAT;%REPEAT;%END

%ROUTINE NONPARM ;! SPEARMANS RANK CORREL
%IF ORD=0 %THENRETURN;NEWLINE;%CYCLE I=-1,1,14;XTAB(-1,I)=0;XTAB(I,-1)=0
%REPEAT;%CYCLE I=0,1,14;%CYCLE J=0,1,14;K=XTAB(I,J);
XTAB(-1,J)=XTAB(-1,J)+K;XTAB(I,-1)=XTAB(I,-1)+K;%REPEAT;%REPEAT
L=0;%CYCLE I=0,1,14;K=XTAB(I,-1)
MRANK(0,I)=(K+L+L+1)/2;L=L+K;%REPEAT;
L=0;%CYCLE I=0,1,14;K=XTAB(-1,I)
MRANK(1,I)=(K+L+L+1)/2;L=L+K;%REPEAT;SIGTX=0;SIGTY=0;PND=L;
%CYCLE I=0,1,14;J=XTAB (I,-1);SIGTX=SIGTX+(J**3)-J
J=XTAB(-1,I);SIGTY=SIGTY+(J**3)-J;%REPEAT;
X=(PND**3)-PND;SIGTX=(X-SIGTX)/12;SIGTY=(X-SIGTY)/12;W=0
%CYCLE I=0,1,14;%CYCLE J=0,1,14;K=XTAB(I,J);%IF K=0 %THEN->L1
V=MRANK(0,I)-MRANK(1,J);W=W+(V*V*K);
L1: %REPEAT;%REPEAT;
R=-750;V=2*SQRT(SIGTX*SIGTY);%IF V<0.000001 %THEN ->L2;
V=(SIGTX+SIGTY-W)/V;%IF V> .99 %OR V< -.99 %THEN ->L2;R=V;SIGGER(PND,0);
L2: PRINTSTRING('SPEARMANS RANK CORRELATION    ');PRCORL;%END
```

```
%ROUTINE ORDERS ;! INPUT COMMANDS
%SWITCH  ORDS(0;20)
L1: %IF ST=1 %THEN NEWLINE;PROMPT('ORDER : ');ORD=RWORD;
%IF ORD >16 %THEN ORD=0;%IF ST=0 %THENSTART;
%IF ORD#1 %AND ORD#6 %AND ORD#7 %THEN ->L1;NEWLINE;%FINISH;
%IF 0<ORD<8 %OR 10<ORD<16 %THEN STARK=1;-> ORDS(ORD)
ORDS(0) : SAY EH;->L1
ORDS(1) : %IF CORLS<13 %THENRETURN;PRINTSTRING('** WARNING :');
          WRITE(CORLS,2);PRINTSTRING(' CORRELATIONS COMPUTED ')
          NEWLINE;PRINTSTRING('** SO');PRINT(CORLS/20,1,1);
          PRINTSTRING(' SIG RESULTS EXPECTED BY CHANCE');%RETURN
ORDS(2) : EVALRO;PRCORL;->L1
ORDS(3) : SPLIT;PRCORL;->L1
ORDS(4) : BIAS;->L1
ORDS(5) : GRAPH;->L1;
ORDS(6) : LIST;->L1
ORDS(7) : FIND;->L1
ORDS(8) : CREATE;->L1
ORDS(9) : RESTRICT;->L1
ORDS(10): CHANGE;->L1
ORDS(11): PARTCOR;->L1
ORDS(12): TTEST;->L1
ORDS(13): ANOVA; ->L1
ORDS(14): ;! DUMMY STATEMENT
ORDS(15): XTABLE;NONPARM;->L1
ORDS(16): ->L1;! DUMMY STATEMENT
%END

%ROUTINE STARTER;! GET PROGRAM RUNNING
PND=0;DEFINE('ST01,OLDDAT');DEFINE('ST02,NEWDAT')
DEFINE('ST03,ENGDAT');STREAM(1)='OLDDAT';STREAM(2)='NEWDAT';
STREAM(3)='ENGDAT';STARK=1;STAR;PRINTSTRING('
ASSORTATIVE MATING IN MAN   PROGRAM MK 11.2
');NEWLINE;ST=3;PROMPT('INPUT : ');OP(1)='+';OP(2)='-';OP(3)='*';
L1: READSYMBOL(Q);%IF Q<'A' %THEN->L1;%IF Q='N' %THEN ST=2;OP(4)='/';
OP(5)='P';%IF Q='O'%THEN ST=1;%IF Q='A' %THEN ST=4;
L2: READSYMBOL(Q);%IF Q#10 %THEN->L2;OP(6)='L';
%IF ST#4 %THENSTART;ALLPAIRS(0);->L3;%FINISH
%CYCLE ST=1,1,3;ALLPAIRS(ST);%REPEAT
PRINTSTRING('THEREFORE  TOTAL INPUT IS '.S.' WITH');WRITE(PND,3);
PRINTSTRING(' PAIRS (');WRITE(CORLS,3);PRINTSTRING(' REVISITED)
');
L3: DSTEP(40)=-10;DSTEP(41)=-10;DSTEP(42)=-10;DSTEP(43)=-10;
CO(-1,1)='?';CO(-1,2)='?';CO(0,1)='L';CO(0,2)='D';CO(1,1)='V'
CO(1,2)='1';FN=0;ST=1;CORLS=0;%END

STARTER;ORDERS;STAR;%ENDOFPROGRAM
```

050133 DART 25K LISTED T40 LP40